DOCUMENT RESUME

ED 050 308 AA 000 704

AUTHOR	Epstein, Marion G.; And Others
TITLE	Selection of Products for Focused Dissemination.
INSTITUTION	Educational Testing Service, Princeton, N.J.
SPONS AGENCY	National Center for Educational Communication (DHEW/OE), Washington, D.C.
REPORT NO	PR-71-8
BUREAU NO	BR-0-0366
PUB DATE	Jun 71
CONTRACT	OEC-0-70-3797(519)
NOTE	140p.
EDRS PRICE	EDRS Price MF-$0.65 HC-$6.58
DESCRIPTORS	*Data Collection, *Information Dissemination, *Instructional Materials, *Standards
IDENTIFIERS	*Educational Testing Service, ETS

ABSTRACT

This task was part of an effort directed to the dissemination and installation of new products. In addition to assembling comprehensive information about these products, the mission was to evolve a set of criteria for use in selecting from among them those to be recommended for National Center for Educational Communication (NCEC) "focused" dissemination attention, to execute a two-step selection procedure, to identify products for extended review and analysis and to recommend products upon which NCEC might focus its dissemination and installation efforts. A nationally representative appraisal panel was convened and met three times over two months to execute this mission. Principal staff support was provided by Educational Testing Service (ETS). Supplementary support was supplied by consultants and staff from the Educational Products Information Exchange (EPIE). ETS staff assembled information principally from developers and also from other conventional sources on all the products in the pool. Criteria were formulated by the panel and on the basis of these criteria, 20 out of 51 products were selected for further study. A review was made of each products's development, and in most cases visits were made to the sites where the developments had occurred. On the basis of further study, nine of the products were recommended for dissemination. For related document, see ED 050 309. (Author/CK)

ED 050 308

PR-71-8

SELECTION OF PRODUCTS FOR FOCUSED DISSEMINATION

Marion G. Epstein
Elizabeth H. Margosches
William B. Schrader
Wesley W. Walton

For the National Center
for Educational Communication
United States Office of Education

AA 000 704

June 1971

EDUCATIONAL TESTING SERVICE
PRINCETON, NEW JERSEY

SELECTION OF PRODUCTS

FOR FOCUSED DISSEMINATION

MARION G. EPSTEIN
ELIZABETH H. MARGOSCHES
WILLIAM B. SCHRADER
WESLEY W. WALTON

A Report on the Development of Product Selection Criteria and of Their Application to the Choice of Available and Verified Products for Nationwide Dissemination and Installation in the Schools

For the National Center for Educational Communication
United States Office of Education

June 1971

Educational Testing Service, Princeton, New Jersey

The task described in this report was performed by Educational Testing Service (ETS) pursuant to Amendment 1 to contract OEC-0-70-3797 (519) with the United States Department of Health, Education, and Welfare, Office of Education, which contract covers operation at ETS of the ERIC Clearinghouse on Tests, Measurement, and Evaluation. The Educational Products Information Exchange (EPIE) assisted in completion of the tasks under a subcontract.

ABSTRACT

This task, completed at the request of the National Center for Educational Communication (NCEC), was part of a new effort directed to the dissemination and installation of new products. In this part of the larger effort, information on verified and available products was requested by NCEC and submitted by principal investigators in the regional educational laboratories, research and development centers, and the colleges and universities. The field was limited in this first year to products that had been developed in projects under the auspices of the National Center for Educational Research and Development (NCERD) and certain other sponsors, all in the U.S. Office of Education.

The mission was to assemble comprehensive information about these products, to evolve a set of criteria for use in selecting from among them those to be recommended for NCEC "focused" dissemination attention, and to execute a two-step selection procedure, to identify products for extended review and analysis and to recommend products upon which NCEC might focus its dissemination and installation efforts.

A nationally representative Appraisal Panel was convened and met three times over two months to execute this mission. Principal staff support was provided by Educational Testing Service (ETS). Supplementary support was supplied by consultants and staff from the Educational Products Information Exchange (EPIE). EPIE staff assembled information principally from developers and also from other conventional sources on all the products in the pool. Criteria were formulated by the Panel in light of this array of product-related data, and reviewed

by NCEC and NCERD staff. Starting with a field of 51 products, the Panel at first applied the criteria using the detailed documentation then available and selected 20 products for further study. Staff analysis on the 20 initial selections included review of the record on each product's development and, in most cases visits to the sites where the developments had occurred. On the basis of its subsequent study of the results of these activities, the Panel then recommended nine of the products for focused dissemination during fiscal year 1971-72.

Foreword

The purpose of this effort has been to provide a foundation of high level and independent judgments by a cross section of the academic and educational communities for new efforts by the Office of Education to disseminate and install new educational products. Typically, panels review and pass judgment upon educational research and development proposals. In this case, it is the result of the execution of proposals, what funded projects "have wrought," that has been the subject of panel evaluation, the result taking the form of new educational products that are available for use outside the developmental setting, and that have already been exposed to satisfactorily concluded field trials.

The task is unique in that it offers the beginnings of a new dimension for program evaluation: a review of an eclectic assemblage of products applicable to the several educational levels from pre-school to adult, for use with diverse populations, covering a broad range of academic areas and cross-disciplines, and treating a wide diversity of educational purposes. It is a review whose objective is to identify from a large product pool a relatively small group of educational products whose widespread dissemination and installation would offer favorable odds for significant improvements in the quality of education in our nation's schools. The ultimate objective is to make product dissemination at once effective and efficient. A large order!

It is important to emphasize that this mission has been most explicit. It has involved the study in varying depths of only limited

segments of the output of the regional educational laboratories, research and development centers, and independently operated projects from which the nominated products have come. The mission has been to assemble information and elicit judgments only about those products under review. Further, it should be made clear at the outset that each product evaluation has been done in the context of the focused dissemination efforts by the U.S. Office of Education, and taking into account the readiness of that product for dissemination at this point in time. As a consequence, the fact that a given product in the original pool does not appear on the recommended list does not itself imply a negative evaluation of the product.

Virtually all the products in this year's pool, regardless of the judgments made about them as part of this effort, should find their way into the pool another year. Some not selected in this cycle would be recommended then, some selected this time might not be selected again.

Finally, neither the descriptions supplied nor the judgments conveyed about a given product should be interpreted as qualitative evaluations either of the center where the product was developed or of the principal investigator and participating staff responsible for its current state.

Wesley W. Walton
Educational Testing Service
Princeton, New Jersey, June 1971

Contents

1.	Nomination of Products for the Pool	1
2.	Establishment of the Panel	7
3.	Initial Collection, Classification, and Organization of Information on Products	15
4.	Mid-Process Additions to the Information on Products	23
5.	Selection Criteria	29
6.	Procedures in Initial and Final Stages of Selection	47
7.	Summary of 1971 Initial and Final Selections* 20 Products Screened for Further Review Nine Products Recommended for Dissemination	57
8.	Recommendations for 1972	63

---

* A separate volume, Products Entered Into the Pool for the Dissemination Program of NCEC, contains non-technical descriptions of products making up the 1971 pool for focused dissemination.

List of Appendices

A PEP Accession List, February 1, 1971

B Product Deferral Lists, January 29 & February 9, 1971

C Product Pool List in Classification Order
(Classification Chart in Text)

D Forms for Project Dossier and Project Log

E Precis Outline and Guidelines for Precis Writers

F Form for Detailed Product Analysis

G Outline of Possible Criteria for Comparative Evaluation of Educational Products

H Formulation of Criteria and Version VI Rating Form

I Revised Formulation of Criteria

J Suggested Procedures for Selection of Products

K Comments on Procedures for Panel Ratings on Products by Michael Scriven and Robert E. Stake

NOMINATION OF PRODUCTS FOR THE POOL

Introduction The pool started with 70 products. By administrative decisions, for the most part related to lack of verification data, 18 products were deferred, leaving 52 in the 1971 Pool for Comparative Evaluation. With one more product removed at the request of its principal investigator, the pool stood at 51. These products were the subject of initial selection efforts. By panel decision, 31 products were set aside, leaving 20 screened products for detailed review and analysis. At the time the panel deliberated over the 20 products, the members debated and finally voted against reinstating one or more of the products that had been set aside. Finally by panel decision, nine products, all among the twenty selected at the first stage, were selected for recommendation to NCEC for dissemination focus. Five were listed as first priority, with four more at a second level of priority.

Origin of the 1971 Product Pool for Comparative Evaluation A request on December 8, 1970 from the Division of Practice Improvement of the National Center for Educational Communication (NCEC) to activities funded by the National Center for Educational Research and Development (NCERD) started the ball rolling. It sought from developers nominations and information on products for entry into the pool, defining products broadly to include "curricula and systems that enhance the learning of students or the operation of educational organizations," and suggesting submissions in cases where products "(a) had been validated by field

testing and (b) were available for widespread implementation at this time." Respondents were requested to describe the products being nominated, using the following outline as a guide:

1. Objectives of the product

2. Population served

3. Product to be installed
   - 3.1 description of the product
   - 3.2 benefits of the product with verification procedures and data
   - 3.3 limitations for a potential adopter (is the product not appropriate for certain groups or does it not work well under some conditions)

4. Cost information
   - 4.1 materials, equipment, and source of availability
     - 4.1.1 required
     - 4.1.2 optional
   - 4.2 facilities
     - 4.1.1 construction
     - 4.1.2 remodeling

5. Personnel
   - 5.1 number and type
   - 5.2 training necessary

6. Administrative considerations
   - 6.1 organizational changes required
   - 6.2 minimum scope of tryout program (individual, class, department, etc.)
   - 6.3 minimum scope of adoption (class, department, school, etc.)
   - 6.4 other

7. Principal investigator's name, address and telephone number

8. If a final report is available giving more detail, please include it

By February 1, 1971, 70 products had been nominated to the initial pool and descriptions on most of them had been received. Appendix A lists the "charter year" submissions.

It was apparent that a number of the nominations, being unavailable for implementation or not far enough along in their verification, had been entered prematurely. Some nominations, moreover, were found to be too divergent from the mission at hand for their inclusion. By mid-February, administrative decisions had been taken by NCEC with the advice and council of Educational Testing Service (ETS) to remove 16 nominations from the pool. (These are listed in Appendix B.) For the most part, the removal action was regarded as a deferral until 1972, or until such time as evidence on the two conditions becomes available.

In the main, the information supplied at the time of nomination by principal investigators and project officers provided an adequate base from which to proceed with product classification and review. The shortcomings that did impose difficulties in the precis-preparation stage had to do mostly with vagueness as to the current state of development and meagerness of validation-related information. Also, it was difficult and occasionally impossible to tell from the information supplied on costs and personnel, what increment of cost the adoption of a given product would impose on a user, above and beyond costs for materials, equipment, and facilities.

In most cases, it would have been highly desirable to have reasonably complete documentation on the product at the same time the basic information arrived on the scene. As it turned out, product review in the typical instance required telephoning the principal investigator for additional documentation, most of which proved to be readily available and easily supplied.

Recommendations for Forming the 1972 Product Pool for Comparative Evaluation Starting in 1972 it would seem desirable that the Product Pool be an ongoing entity, a way of queueing products as they are nominated by their developers as ready for national use. As soon as a product is thought to be at that state of development, NCERD and NCEC could determine whether or not it should be subjected to comparative evaluation. Those products receiving affirmative decisions would go into the Product Pool scheduled for the next comparative evaluation, and information on them would then be collected as a matter of course, either by NCEC or its contractor.

Further improvement would be gained if guidelines to nominators were aligned with the criteria against which the 1971 Panel made its selections. For example, in regarding a given product, the panel members leaned heavily on developers' definitions of product goals and objectives, and evidence of effectiveness within the context of these statements of purpose. It would be constructive in the future to invite developers to describe objectives and verifications side by side, and to express in guidelines that the latter be viewed as an explication of achievements related to the former. It would also be advisable to separate data used in formative evaluation from data that are summative in nature. Another example of closer alignment of guidelines to criteria is in the area of costs. The panel sought information on "non-dollar" costs beyond expenditures for materials, equipment and facilities; for example, equivalent costs of training time, substitute costs during in-service training periods, costs for monitoring quality, and so on. Cost detail should be extended where

feasible to include estimates of cost for the implementation of all the support systems essential to make the product operational in the settings for which it was developed.

The above illustrations are cited to suggest that the request for information on nominated products reflect as fully as possible the most current statement on criteria for use in the Product Evaluation Project. It would also be highly desirable for the request for product nominations and information to elicit as complete a documentation on a nominated product as its principal investigator can supply.

Now that the Product Evaluation Project (PEP) has had the benefit of an initial cycle, it would seem desirable to call for nominated products in such a way as to deter submission of marginal candidates. Introduction of a screen at point of origin coarse enough to foster the nomination of promising products that meet general needs and fine enough to deter nomination of obvious misfits might be a step that would move the project forward in a positive direction. In the absence of such a screen, the art of comparative evaluation will be in for a hard time. One senses that perhaps as few as 25 of the products in the initial 1971 pool might have passed through such a screen, and in 1972 without preselection there might be as many as 100. A reasonable and realistic expectation for 1972 in light of the recent experience would not exceed the range of 25-30 - unless of course productivity in educational research and development suddenly accelerates by a substantial degree. If these estimates are right, the preselection screening suggested would foster approximately one out of four in an open field.

At the same time, it would seem desirable to seek out promising products from a larger universe, including, for example, materials from projects under the auspices of the National Science Foundation, and the Office of Economic Opportunity.

A final comment is in order related to the question of timing. Whether ongoing or on a deadline, the product nomination process needs an earlier start in order to execute an optimum time table. The critical need is to stretch out the schedule so that panel meeting I on the criterion question (which will utilize some by then available product descriptions and precis) can be held early with six to eight weeks between that first meeting and the Panel's meeting II, for its initial phase product selection. Better staff preparations then could be made, closer account could be taken of the panel's criteria, and the product precis could be made more useful to panel members at both initial and final selection meetings.

ESTABLISHMENT OF THE PANEL

Introduction Critical to the success and long term viability of such a project as this is acceptance of responsibility for key decisions by a top-flight group of disinterested parties representative of change agents, curriculum planners, and materials-users. A first order of business in this project, then, was the selection and commissioning of an Appraisal Panel.

The Appraisal Panel of eight to ten people was to include curriculum workers, evaluators, teachers of teachers, and measurement specialists, some of whom were to be members of special target populations for which new products had been developed. It was also to provide links to the sternest critics of educational research and development and to those who participated in establishing the program of OE-sponsored educational research.

Educators who had been involved during the formative period of government-sponsored educational R & D were sought as Panel members to have some way to determine whether the products and materials now coming from these concentrations of intellectual energy are realizations of the early-day expectations. The severest critics were tapped to assure credibility for the evaluations done.

The panel was to be the decision-reaching body. Other individuals related to the project were to be viewed as sources of "staff support" to make the panel's tasks in the course of reaching decisions as effective and pleasant as possible. The panel's major missions were (1) establishing criteria for selection of products in two stages,

(2) at the initial stage, setting, reviewing, revising, and confirming a list of selected products for study in depth, and a list of products to be set aside for further development, and (3) at the second stage, selecting from among the initially chosen products, after they had been exposed to detailed analysis, those products to be recommended for special focused dissemination.

Product Evaluation Panel, 1971 A panel of eight was decided upon, and it was regarded as essential that members should provide direct representation from among the following sectors:

- o Curriculum and Instruction
- o Teacher Training
- o Evaluation
- o Product Developers
- o School Administration
- o Philosophy and History
- o Urban Affairs
- o Lay Critics of the Public Schools

Clearly, in covering so many fields with so few, it would be necessary to identify potential panelists whose credentials in most cases would span several of the sectors sought. As it turned out, there were 24 affiliations with the sought sectors among the 8 panelists appointed. In effect, each panelist covered three fields. At the same time reasonably good geographical distribution was achieved.

Moreover, it was the project's good fortune to have among the panelists two women, a man sensitive to minority group problems, and

a man currently in education with previous wide experience in commercial publishing and distribution.

Members of the 1971 Product Evaluation Panel, together with summaries of credentials are listed below. Dean David R. Krathwohl* served as the panel's chairman.

PRODUCT EVALUATION PANEL

1971

Richard Gousha	Superintendent of Schools, Milwaukee, Wisconsin 1967-
Formerly	Superintendent of Public Instruction, Delaware 1964-67 Superintendent of Schools, Cuyahoga Falls, Ohio 1960-64 Assistant to Director, Research and Field Services, University of Indiana School of Education 1959 Superintendent of Schools, Amherst, Ohio, 1956-59 Teacher and local school executive 1947-56 (several places)
Other relevant experience	Advisory Committee Milwaukee Children's Court American Research Council Great Cities Program for School Improvement Board of Directors, Milwaukee Technical College, Milwaukee Symphony Wisconsin Regional Board, National Conference of Christians & Jews
Education	A.B. Heidelberg College (Ohio), M.A. Western Reserve, Ed.D Indiana
Robert Heinich	Professor of Education, Division of Educational Media, Indiana University
Formerly	Head, Department of Instructional Systems, Doubleday Publishing Co.
Education	Ph.D Instructional Technology - University of Southern California

*David R. Krathwohl	Dean, School of Education, Syracuse University 1965-
Formerly	Director, Bureau of Educational Research, Michigan State University 1963-65 At Michigan State since 1955. Research Coordinator, and Professor, Bureau of Education Research Assistant Director, Unit of Evaluation, Bureau of Educational Research, University of Illinois 1949-55
Other relevant experience	President, American Educational Research Association 1968 Chairman, Board of Trustees Eastern Regional Institute for Education
Fields	Educational measurement, problem solving processes, school learning - co-author Taxonomy of Educational Objectives, The Classification of Educational Goals - a major contribution
Education	B.S., M.A., Ph.D University of Chicago
Ruth Mancuso	Member, N.J. State Board of Education - Chairman, Committee on Evaluation Director, Audiovisual Center for Tri-County Area in South Jersey
Formerly	President, National School Boards Association President, New Jersey State School Board Association President, Local School Board, Glassboro, New Jersey Chairman, N.J. State Committee on School Regionalization Member, N.J. State Committee on Vocational Education Public School Teacher
Other current activities	National Assessment Advisory Panel Advisory Committee for ERIC Clearinghouse on Tests & Measurements

Mary Molyneaux	Assistant Superintendent, Curriculum and Instruction, Pittsburgh, Pa. Public Schools 1965-
Formerly	Director of Curriculum, Pittsburgh, Pa. 1964-65 Principal, 1945-64 Supervisor of Elementary Education 1940-45 Lecturer, University of Pittsburgh 1945-58
Other relevant experience	Yearbook Committee, Department of Elementary School Principals, NEA Steering Committee, Curriculum Continuity Demonstration, University of Pittsburgh Committee on Instructional Materials, Research Council of the Great Cities Programs for School Improvement
Education	B.S., M.A., Ed.D: University of Pittsburgh
Michael Scriven	Professor, Philosophy Department, University of California at Berkeley 1966- Presently on leave as Whitehead Fellow at the Harvard School of Education
Formerly	Professor of History and Philosophy of Science, Indiana University, 1960-66 Fellow, Center for Advanced Study in Behavioral Sciences 1963 Taught at Swarthmore College, University of Minnesota 1952-60
Other relevant experience	Editorial Boards of Encyclopedia of Philosophy, American Philosophical Quarterly, Metaphilosophy, Contemporary Psychology and editorial consultant for Science. Board of Directors, Social Science Education, Inc. 1963- Director, Evaluation of Education Materials Project, SSEC, 1963-67 Director, SSEC project on role of values in the social studies 1963-66 Chairman, Advisory Board for Evaluation, Central Midwestern Region Educational Laboratory, Inc. Advisory Board, Social Studies Program, Minnesota Mining and Manufacturing Corporation, Project Follow Through 1968-69 Consultant, U.S. Office of Education 1968-69 Evaluation Consultant, Marin School Board, Social Studies Project, 1969-
Fields	General philosophy, philosophy of science, psychology, psychiatry, education, logic, computer technology
Education	A.B., M.A. University of Melbourne, Ph.D Oxford

Robert E. Stake	Professor, Educational Psychology, University of Illinois 1966-
Formerly	Assistant, Associate Professor, University of Nebraska 1958-66
Fields	Psychometric methodology, programmed instruction, use of computers for educational research
Education	B.A., M.A.: Nebraska, Ph.D: Princeton
Bernard Watson	Professor and Chairman of Urban Education, Temple University (developing new program) 1970-
Formerly	Deputy Superintendent for Planning, Philadelphia, Pa. School District Associate Superintendent for Innovative Programs, Philadelphia Staff associate for midwestern administrative center at University of Chicago Teacher, Counselor, Vice-Principal and Principal in Gary, Indiana
Other relevant experience	Local Boards for Urban Coalition and Model City Programs Board of trustees for two private schools A variety of other committees for such things as National Teacher Corps, Ford and Rockefeller Foundations, etc. Lectured at Princeton, Yale, and University of Pennsylvania
Education	B.S.: Indiana University, M.A.: University of Illinois, Ph.D: University of Chicago

Product Evaluation Panel, 1972 The criterion for the composition of future panels would be to match or beat the one functioning in 1971. Admixtures of background could not have been better. Rapport established at the outset brought issues squarely onto the table. Open, sincere and effective arguing brought early incisive perceptions that seemed to make decisions down-stream easier to reach. A panel as diverse as this, obviously, did not always achieve consensus. It did, nonetheless, find ways of reflecting individual views in concert.

In its own reprise, the panel emphasized the advantages of plenary action, and the strengths it drew from common exposure to the domain within which its decisions were needed. One might surmise that it would not have looked as favorably on an option that would have broken the group into two or three more operationally convenient working panels.

In 1972 and thereafter, it would be desirable to increase panel size to ten. Moreover, there should likely be an overlap of at least two, and not more than three, members in the panels for successive years. For example, the panel needed a researcher practicing in the public school system and a practicing elementary or secondary classroom teacher.

INITIAL COLLECTION, CLASSIFICATION, AND ORGANIZATION OF INFORMATION ON PRODUCTS

Introduction The problem in bringing usable information together on this widely diverse pool of products was to devise ways of highlighting similarities and differences. Each product had been selected with care by the funded organization from a multitude of possible products it might have developed. In most instances, the center had been aided in making the decision on what to develop by a distinguished group of educators serving as advisers. Moreover, the products had been developed with care by professionally competent staff. From the nature of the materials developed, one could surmise that the staff had used its best judgment in seeking advice from others and in conducting field trials in the formative stages of product development.

Beyond these two common characteristics, however, there were a number of differences. First, from the standpoint of national need, the products could be arranged along a continuum. Some products were directed toward educational problems of critical importance; others addressed themselves to instructional or administrative situations of lower current priority.

The products differed too in the extent to which their use in the field had demonstrated their effectiveness. The value of some newly developed materials, for example, had not yet been demonstrated. Extensive verification under a variety of circumstances was yet to come. On the other hand, some products had been shown to be effective in varying degrees.

There were also differences in the degree to which a given product was more effective than existing materials or methods. If the new materials or procedures had produced results which were significantly superior to products already available, this fact needed to be known and taken into consideration in the selection of products for special dissemination attention.

Finally, variations among products were observable in terms of implementation cost and effort. The introduction and maintenance of some products would require extensive staff training, the purchase of expensive materials or equipment, and even drastic revision in the use of available people, equipment, and space. Other products required minimal outlays of this nature.

Classification and Organization of Information To immediately identify similarities among products, each product description on receipt was read and a four-digit classification number was assigned to categorize that product along four dimensions as follows:

X000 Product Purpose

0X00 Target Population

00X0 Educational Level

000X Academic Field

Figure 1 shows the Product Classification for 1971. A listing of the 1971 Product Pool in order by purpose, target, level and field appears in Appendix C.

A second step in materials organization was the compilation of all developmental, evaluative, and other descriptive information on

1971 Product Pool

Products entered in the 1971 Product Pool for comparative evaluation have been categorized in ways that will allow comparison of the material. Each product has been classified along dimensions as noted in the four columns below. A four-digit code, then, describes for a given product its major thrust, target population, educational level, and academic field.

1971 Pool	Product Mostly for	Target Population	Educational Level	Academic Field
7	1 Ending Neglect	X1 Blacks	XX1 Pre-School & Early School	XXX1 Language Arts
7	2 Individualized Instruction	X2 Mexican-Latin Americans, Indians, Eskimos	XX2 Middle School	XXX2 Foreign Language
12	3 Group Instruction	X3 Disadvantaged Whites	XX3 Junior & Senior High School	XXX3 Science
15	4 Teacher Training	X4 All Disadvantaged Groups	XX4 College	XXX4 Mathematics
7	5 Administering Programs	X5 Disadvantaged and Other Learners	XX5 Cross Level	XXX5 Social Studies
3	6 Two or More Programs	X6 Academically Talented	XX6 No Level	XXX6 Humanities
	7 Something Else	X7 Adults		XXX7 Occupational/ Technical
	0 Undetermined	X8 Students in a Field		XXX8 Fine & Performing Arts
51		X9 Common Interest		XXX9 Multi-Disciplinary Cross-Field
		XX No Target		XXX0 No Field (A method, technique, system, or classroom management)

Figure 1 - Product Classification

each product into a dossier. To the dossier was added additional information assembled during the course of the review, the results of detailed analyses, records of Appraisal Panel judgments, and the like. The dossiers, then, contain records of both information and actions regarding every product, and will be retained as an official ERIC-TM resource.

The ongoing collection of information for each dossier was the responsibility of a two-man product review team who also had the task of drafting a product precis, which included the most important of the panel's evaluation needs. The purposes of the precis are to bring each product to a common base for comparison with the others, to highlight differences among products, to expedite study and review by panel members, ETS specialists and OE representatives, and to compile descriptive information for later use in dissemination. The precis is the basic document for each dossier.

Precis writers were drawn from among experienced members of the professional staff in ETS's Test Development Division. Those working on a given product, typically, were examiners in that discipline into which the product would most naturally fit. Wide field contact in their respective disciplines meant that in many cases staff members already had known of the development from which the product emanated as well as something about competing developments.

In order to assure that the precis were sufficiently communicative in conveying accurate and adequate information about a given product, these staff write-ups were sent to principal investigators for their review and comment. Responses from the developers were then incorporated

into a precis revision which was put into the documentation for the panel's use.

A check list serving as the front cover for the dossier (blue) and a similar record (pink) in a "log book" were used as instruments to monitor the generation and completion of the collection in each dossier. These forms, revised to incorporate the 1971 experience, appear in Appendix D. In addition to the control documents, a summary listing of contents was kept on the dossier's back cover, and a wall chart provided an up-to-date graphic picture as progress was made in acquisition and arrangement of product-related materials and in preparation of precis.

These kinds of management aids were essential to the success of this operation, which was on such a tight time schedule and involved staff support by so many people. A large portion of this task was concentrated into a seven week period of work. The scope of staff involvement may be measured by the size of a distribution list for a note of appreciation for substantive help in meeting the schedule: It went to 30 professional staff members of ETS and EPIE.

As noted above, the precis was the key document for panel consideration of a given product. It was to function as a succinct description, and also was expected to flag sources of additional detail in the dossier. Through the precis the panelist was to be able to assimilate what was in the record on a product with minimum efforts. The four page form designed to channel efforts of precis writers and the set of guidelines issued as a further aid in precis preparation are shown in Appendix E.

Indexed loose leaf books, one for each panelist, were prepared as the working documents for selection meetings. Each contained all the precis on products due for panel consideration. The order within which product precis were entered into the books was set by the classification scheme, specifically:

- Ascending order in each class in turn
- Randomly for products with the same ordinal position in a given class

The major advantage of this arrangement was to place products before the panel for deliberation in compatible clusters. For example, products designed for problems of neglect were dealt with separately from those related more directly to teacher training. At the same time, the effect was to randomize the products so that consideration of a product was unaffected by its position in the book.

The remaining pieces of information to which the panel could refer at will and without inconvenience were lodged in the dossiers on all products. Thus, the entire available record, including ERIC reports in microfiche, was accessible to panelists and received extensive use during periods of panel product review.

Needed Changes in Presentation of Information The classification scheme needs changing so that it will serve to describe products more fully. Its expansion will make it possible to assign to a product a number of descriptions out of each category. Depth indexing after the fashion of ERIC Clearinghouse methods might be a more satisfactory model than the four-dimensional classification as used in 1971. In the future,

the time table will be eased considerably, especially during the early steps, as described above. The most needed schedule adjustment is to increase the span of time between nomination of products for the pool and completion of the precis for the panel's use. With more time, precis could be prepared by a much smaller group of product reviewers, thus assuring consistent and consistently high quality precis, and reducing concerns about inter-precis reliability.

Precis-writing calls for a special set of high level skills in the acquisition of which training plays an important part. In the future, a team of not more than 10 product reviewers, adequately oriented and trained in the special skills of precis preparation, should share the responsibility of learning what is known about all the nominated products and communicating it to the panel members.

Improvement in overall precis quality, which should be an objective for the second cycle in any case, requires both careful selection of product reviewers for their general verbal facility and the furbishing of their skills for the specific mission at hand.

Panel members have suggested the following framework for the training task:

a. A statement of product-selection criteria to be used by the panel, or in its absence the statement of criteria used by the previous panel. This should serve to set the stage.

b. Good examples of precis and poor examples, together with supporting documentation on the products described, might provide the models within which training activities could be organized, after the fashion of the Minicourses, AC28 and AC42, found elsewhere in this report.

An exercise is currently in progress to help advance these ends, resulting from an end-of-session action by the panel. Three to five precis, thought by the ETS staff to be of good quality, are being sent on the mail circuit, a set to each panel member. Each panelist will rank each precis on a scale of A to F and return his votes. Staff members, then, will have in hand a 1971 panel representation of precis quality for use in setting quality standards for 1972.

The changed initial product review procedures discussed above in combination with the supervisory review and editing procedures used this year should satisfy the needs in the area of "information and data reduction."

A change is also in order in the manner of organizing the loose-leaf book of product precis for panel review. Though it might appear an insignificant detail, its importance cannot be overemphasized. Once the book order of product precis is established, each precis should be assigned a book sequence number, and thereafter be referred to by that number. The number, preferably a gummed label, should be affixed to the upper right hand cover of precis page 1. An index showing sequence numbers and product names should be inserted as a frontispiece. These improvements will make it easier for panelists and staff members to find a given product precis on short notice.

MID-PROCESS ADDITIONS TO THE INFORMATION ON PRODUCTS

Introduction A preliminary selection from the product pool resulted in the identification of twenty products for final 1971 evaluation. It was necessary to secure the most up-to-the-minute and comprehensive information about these products and their development to assure that final panel judgments could be made with all pertinent pieces of the picture in place. Completion of this part of the task was under severe time constraints; there were 34 calendar days between the preliminary selection and the final selection sessions of the panel. During the interim, each product was surveyed by a product-evaluation specialist, reviewed by a subject-matter specialist, and in the case of seventeen out of the twenty, site visited by a generalist. The results of these analyses were made part of the array of product-related information used by the Panel as it made its final selections.

It was initially intended that summary information, both quantitative and qualitative in nature, be assembled at this stage. The former was to be an attempt to reduce evaluative observations and data to numerical terms. The approach the panel took to solving the evaluation problem - largely impressionistic, highly judgmental - cancelled earlier plans which incorporated quantitative methods. The augmentation of product-related information between the initial and final stages by generalists, and by subject-matter and evaluation specialists, focused upon only criterion-based considerations since by then it was known that these would be the factors panel members would take into account at the final selection meeting.

Additions to Information on Products The analysis on the 20 products initially selected was in two stages. During the two days following the panel's initial selection meeting, a task force of three product evaluation specialists from the University of Illinois, functioning under ETS's subcontract with the Educational Products Information Exchange (EPIE), exposed all the available documentation on the selected products to painstaking review. The task force's mission was to identify missing elements in the documentation and to suggest additional information that would be particularly helpful to have during the consideration of these products at the final selection stage. Setting to work, this group produced a written critical review of each of the 20 products, highlighting shortcomings on documentation and noting apparent flaws in research design. These analyses were supplied to the product reviewers and site visitors responsible for follow-on activities related to the respective products. Since this task force's work products were means to ends, rather than ends in themselves, the output of this stage is not a part of the record.

The second stage consisted of concurrent schedules for completion of (a) detailed product reviews and analyses, and (b) site visits. There was a different reviewer for each product who, in almost all cases, was a subject-matter specialist with at least some prior knowledge of the development put under his purview. Site visits were made by means of circuit-type itineraries, mainly to economize on travel expenses. Typical circuits involved two or three site visits. One traveler visited six developers at four sites.

The work product from both product reviewers and site visitors was the same: a detailed product analysis "report" to be entered into the record as a supplement to the product precis. Reporters were asked to organize their detailed product analyses around three information clusters as follows:

- o Goals, effectiveness, and the relationships among them
- o Costs, implementation questions, and the relationships among them
- o Open-ended comments including conclusions and judgments.

The form that guided preparation of the precis augmentations appears as Appendix F. Orientations to the tasks were accomplished in individual conferences.

There was one product among those selected by the panel for study in depth that was dealt with in a unique way. Special action was taken in this instance to avoid a possible conflict of interest on ETS's account. Before the Product Evaluation Project was begun, the developer had approached ETS to assume reponsibility as publisher and further developer of his product, and ETS at the time had under development a product similar both in purpose and execution. The entire staff review and analysis of this product was conducted by EPIE, whose consultant reported his finding directly to the panel without intervention by ETS.

In any case, reports by product reviewer and site visitor were added to the precis on each of the 20 products. In some instances, the precis themselves were revised to reflect the most current information supplied to the project office by the principal investigator.

For its final selection meeting, the panel had blue loose-leaf books containing precis and detailed product analyses on the twenty

products selected during the initial stage. Each panelist also had his earlier black loose-leaf book, now containing precis on the 31 products set aside, for which some precis had been brought up to date since last meeting.

It seems reasonable to state that information available to the panel at its final selection was considerably better than at the initial stage. Significantly improved product descriptions seem to have come from the in-depth exposure which detailed analysis and site visits afforded.

Refinements on Supplementing Information on Products Excepc for the limitations of time, this part of the procedure seemed to work very well. Had more time been available, it would have been productive to articulate into a single document (a) findings from detailed analysis, (b) results of site visit, and (c) the latest changes in basic product-related information. Ideally, this would take the form of a final and comprehensive version of the precis with additions clearly identified. An all-inclusive document such as this would make for a more unified approach by the panel and a somewhat simpler context for the panelists to cope with in dealing with complex questions of judgment.

This pattern for increasing the quality of product-related information suggests that site visits should be made by the precis writers. Were the same individual to see the information gathering and analysis tasks through to their completion, he would become an authoritative resource.

The panel's dominant view was that the site visits added a great deal to the body of information available, but that to have a pre-

structured interview guide for use by all site visitors would be a good idea. Some members questioned whether actual observations of products in use could be expected to produce much in the way of useful information, though there was general agreement that direct discussion with principal investigators and staff members should be encouraged.

The panel members took the opportunity to query both product reviewers and site visitors for clarification and further explication of their findings. Having these "expert witnesses" readily at hand seemed to expedite the proceedings. It would be advantageous in the future to have them automatically on call should their "testimony" be needed to advance the panel's consideration of a given product.

SELECTION CRITERIA

Introduction The criteria developed by the panel were related to what the panel conceived as its primary task -- to identify products which it could recommend to USOE for dissemination in 1971. The narrow focus of the task served to sharpen the definition of criteria. It was recognized that both a product and a strategy for its dissemination could change significantly in a year. The criteria, accordingly, should be interpreted in the light of the immediate purpose they were intended to serve.

Although the number of major criteria is relatively small, each includes a number of more or less disparate elements. The grouping of a large number of characteristics under a few major headings for purposes of rating presumably made rating decisions somewhat more complex but helped to organize the evaluation process. The major headings used for grouping criterion-related elements were: goals, effectiveness, "costs," and adoptability. The detailed judgments made in the course of reviewing a product in terms of specific criteria served as the basis for making an overall judgment about the product. In making his global judgment the rater took account also of USOE priorities and other supplementary considerations.

The panel devoted most of its first meeting on March 3 and 4, 1971 to the formulation of a detailed set of criteria and to devising a plan for using these criteria in evaluating products. Professor Scriven designed a series of rating forms reflecting the evolution of the panel's conception of the criteria and of the rating scales and

checklists for recording judgments about a product. A copy of the seventh version of the rating form appears as Figure 2. Because this discussion of the criterion is closely related to the rating form, the reader should find it useful to refer to the form while considering the description of the criteria. Although the rating form served a number of purposes, it was particularly useful as a worksheet for the rater both when making his overall rating and during subsequent discussions of the product.

Criteria for Selection of Products in 1971

Goals of the Product's Development: Extent to which product may be expected to have major effects on significant educational outcomes. Four separate aspects of this criterion were considered:

(1) Urgent Present Need - Does the product address itself to urgent needs?

(2) Desirable Originality - Does the product embody well-conceived innovations in content, method, or both? Does it reflect old orientations or new ones?

(3) Educational Centrality - Does the product concern itself with outcomes that are central to education rather than with special outcomes?

(4) Size of Target Population - How large is the group for whom the developer considers the product appropriate?

Raters designated any of the four aspects on which a product was judged to be notably strong or notably deficient. If some other aspect of the product's goals deserved comment, the rater could

RATING FORM – PRODUCT EVALUATION POOL '71 – ETS

--------DRAFT VII----------

PRODUCT CODE NUMBER: AC _____ (Further descriptions if desired; Aim, or Author, Institutional Sponsor etc. .....

RATER'S NAME / CODE: ..... Date of Rating / /71

GOALS [A = Excellent B = Good C = Possible D = Unimpressive]

Considerations: Urgent Present Need / Desirable Originality / Educational Centrality / Size of Target Population / ..... (other)

Circle those of unusual positive significance.
Strike out any that struck you as notably deficient.

Rating: A B C D (Circle one or more; or add ± if preferred)

Indicate here if alternative conception of goals or target etc. has been used as a basis for second rating sheet ☐

EFFECTIVENESS [A, B, C, D, or "F" = Undesirable, or "?" = No reasonable estimate possible at this time]

Using data of 1971

Considerations: Adequacy of Test Data (Sample Size; Sample Fairness; Appropriate Controls; Objectivity of Judgment) / Performance in Field Trials Including Teacher Feed-back / Background or Internal Evidence Suggesting Success or Failure e.g. Prior Track Record

Rating on Stated Goals: A B C D F ?

Rating on Side Effects or Unanticipated Effects: A B C D F ? ..... (Describe Side Effect)

"COSTS"

To User Considerations: Extra (or less) Professional Staff / Extra (or less) Maintenance Staff / In-service Training Costs / Consultant Costs / Repair and Substitute Costs / Extra Space Costs / Extra (or less) Time Required / System Disruption / Opposition ·Students ·Staff ·Community

Rating:

Costs of Installation:	Negligible	Modest	High	?	COSTS SAVINGS
Costs of Maintenance:	Normal Range	High	Very High	?	
Non-dollar Costs:	Minimal	Tolerable	Excessive	?	

ADOPTABILITY

MATERIALS: Immediate Unlimited Availability / Immediate Limited Availability / Sample Materials Only / Evaluation Materials Availability

ADMINISTRATION: Requires Special Training in Advance of Use / Likely to be System-disruptive / Plant Installation Required (Lead-time.....years)

Rating: Easy A B Few Problems C D Too Difficult F ?

OVERALL

Considerations — Circle factors, if any, that particularly affected your judgment

Above Judgments About: Goals Effectiveness / "Costs" / Adoptability / USOE Priorities / Other Support Available (or not available) / "Impact" or "Leverage" or "Multiplication" Effect of Support / Existence of Alternative Products comparable in adoptability and at least equally cost-effective and desirable

[CRITICALLY COMPETITIVE PRODUCT(S) .....]

Rating of need for USOE 1971 dissemination support based on data available 4/71: A B C D F

Figure 2 - Rating Form VII

describe this aspect briefly and designate whether it was a strong or weak point of the product. The summary rating for this criterion was expressed on the following scale:

A Excellent

B Good

C Possible

D Unimpressive

A plus or minus sign was added to the rating if the rater wished to do so.

Provision was made for the possibility that the rater would wish to evaluate the product under different assumptions concerning the product's goals. If so, he was asked to show his alternative evaluation on a separate rating sheet. This option was useful, for example, if the author's claimed goals for the product were inconsistent with the product as developed, in the judgment of the rater.

The panel reached consensus on this criterion with little discussion except to clarify the meaning of "Educational Centrality." In rating this characteristic, it was decided to take the author's claims at face value even though other evidence suggested that his goals might not correspond exactly to the outcomes produced.

Effectiveness of the Product: Extent to which the product is effective in accomplishing its stated goals in its target population and in accomplishing goals other than those stated by its developer or producing outcomes in populations other than its target population (side effects). Judgments on this criterion were based on evidence available in the spring of 1971. Three aspects of effectiveness were considered:

(1) Adequacy of Test Data -- How adequate are test data with respect to sample size, fairness of sample with respect to target population, provision of data for appropriate control groups, and objectivity of judgments about the product?

(2) Performance in Field Trials -- How well did the product perform in field trials? Is there evidence that teachers and students accepted the product readily and wish to continue using it?

(3) Internal and Background Evidence -- Is there internal evidence of product quality or evidence of its background (for example, previous outstanding performance of the developers in producing highly effective products) which offers useful supplementary indications of the product's probable success or failure? Is product content appropriate to stated goals?

With respect to side effects, the rater stated briefly the alternative goals, alternative populations, or unanticipated outcomes with which his rating was concerned. For example, a programmed text in algebra might have produced gains in reading ability. On the other hand, adverse emotional effects might arise from certain ways of teaching reading. The observations of the raters with respect to side effects were used to seek further evidence when site visits were made.

Separate ratings on effectiveness were made with respect both to stated goals and side effects. For stated goals, the following scale was used:

A Excellent

B Good

C Fair

D Unimpressive

F Undesirable

? Impossible to estimate on the basis of existing evidence

For side effects, certain letter ratings were modified as follows:

C Neutral

D Somewhat negative

F Undesirable

Here again, a plus or minus sign was added to either rating if the rater wished to do so.

The development of this criterion entailed several decisions by the panel. First, the panel rejected the idea of a separate rating on "educational soundness." Instead, internal evidence of effectiveness of a product was considered along with adequacy of test data and performance of the product in field trials in assessing effectiveness. Evidence obtained through empirical studies and actual use in schools thus were treated as essential to the evaluation of product effectiveness. Second, a good deal of thought was devoted to a proper handling of "side effects" -- that is, positive or negative effects produced by the product other than those intended by its author. The decision to permit a separate rating of effectiveness based on side effects enabled the rater to evaluate effectiveness both in terms of the author's goals and in terms of the way the product was actually operating.

Costs: Extent to which the introduction and subsequent use of the product place heavy demands on the school's resources, both economic and human. In evaluating the costs of a product, the rater was asked to consider both financial outlay and a cluster of 10 other aspects as follows:

(1) Materials (inadvertently left out in the printing of Draft VII)

(2) Increase or decrease in professional and/or paraprofessional staff

(3) Increase or decrease in technical staff (audio-visual, computer, etc.)

(4) Inservice training costs in dollars and time

(5) Consultant costs related to adoption and maintenance

(6) Repair and substitute costs

(7) Costs for extra space

(8) Increase or decrease in time needed by students, teachers, administrators

(9) Disruption of the system

(10) Opposition by community, students, and staff.

The last three factors received special consideration since products which can easily be introduced without disturbing the ongoing system and products which can be introduced without extensive involvement of higher-level administrators are particularly likely to repay dissemination efforts. On the other hand, dissemination of products which, for example, seem to teachers to downgrade their professional role, might be expected to encounter serious resistance.

Separate ratings on costs were made with respect to installation, maintenance, and non-dollar costs. For installation, the three levels were: Negligible, Modest, and High. For maintenance, the three levels, Normal Range, High, and Very High, were related to annual per-pupil costs for the more traditional ways of achieving the same objectives. The following estimates for the annual maintenance cost of a single course were used as guidelines: Normal Range ($0 to $4), High ($4 to $14), and Very High (More than $14). For non-dollar costs, the three levels were Minimal, Tolerable, Excessive. Raters could make a separate rating of non-dollar costs or they could take account of non-dollar costs in their ratings on installation and maintenance. For total installation, maintenance and non-dollar costs, the rater indicated uncertainty by using "?" as his rating. Space was provided to record the rater's judgment of estimated costs or savings.

As might be expected, the evaluation of costs presented a number of problems, despite a consensus that cost was indeed a significant factor in judging the desirability of disseminating a product. The difficulties of comparing dollar figures for products having different scope of applicability and of deciding what assumptions to make concerning equipment already available to the school and concerning the number of pupils and of years over which costs might be distributed were recognized. It was decided that it should be feasible to compare the costs of a product with conventional means of achieving similar outcomes, and it turned out that judgments of this kind could be made. Another problem arose in deciding whether the noneconomic costs of a product (e.g., system disruption) should be converted roughly into

economic terms and added to the economic costs or should be handled as a separate variable. It was decided to let the handling of this point be optional with each panel member.

Adoptability: Extent to which the product is readily available to a school which wishes to adopt it. Four considerations related to materials were specified as follows:

(1) Immediate, virtually unlimited, availability

(2) Immediate availability, but supply is limited

(3) Only sample materials currently available

(4) Materials for use in product testing and evaluation available

Three significant administrative considerations affecting adoptability also were specified, as follows:

(1) Requires special training in advance of use

(2) Likely to be system-disruptive

(3) Plant installation required

Space was provided to indicate lead-time, in years, for installation.

Adoptability was rated on the following scale:

A Easily adoptable

B

C Few problems

D

F Adoption too difficult

? Impossible to estimate on the basis of existing evidence

In the course of the panel's work, the "adoptability" criterion developed from a simple check list concerning availability and implementation to a full-fledged rating variable. This change probably resulted from an increasing recognition of important differences among products with respect to this characteristic. Both the current status of the product and certain basic characteristics were regarded as significant in making this judgment.

Overall Judgment of Need for USOE 1971 Dissemination Support: This was a summary judgment of the extent to which the product deserved favorable consideration by the Panel for inclusion in the recommended group of products. Judgment was based on the most up-to-the-minute data available, and was to encompass in one global estimate the separate judgments made regarding goals, effectiveness, "costs," and adoptability.

In addition to epitomizing the specific criterion-based evaluations, the overall rating also took account of the following:

(1) Concurrence of product with USOE priorities, as follows:

A. Previous priorities

- o Right to Read
- o Disadvantaged - elementary and secondary
- o Equal Educational Opportunity
- o Enviromental/Ecological Education
- o National Institute of Education
- o Experimental Schools
- o Disadvantaged - postsecondary

B. Present priorities

- o Create opportunities through education of the handicapped
- o Stimulate development of innovative and effective approaches to education
- o Meet the needs of economically disadvantaged children
- o Eliminate racial, ethnic and cultural barriers to educational opportunities
- o Stimulate career education programs

(2) Possible effects of NCEC dissemination efforts considered in the light of existing support (if any) for dissemination. Are sufficient dissemination efforts already being made? Would this dissemination support provide needed impetus to get the product over the hump?

(3) Impact or leverage or multiplication effect of support. This includes size of market, visibility of results, and breadth of goals sought. Would dissemination support accelerate implementation to a significant extent? Would it accelerate school improvements?

(4) Existence of alternative products, comparable in adoptability and at least equally cost-effective and desirable. Does this product introduce something unique? Are there equally good products to serve the same need as effectively and in the same range of "costs"? Space was provided for designating critically competitive products.

The rater made notations on those aspects of which he had taken account, identified critically competitive products, and then subsumed all the component judgments into a final rating.

Ratings for the global criterion were expressed on a five-step scale as follows:

A Excellent

B Good

C Fair

D Unimpressive

F Undesirable

A plus or minus sign was added to the rating if the rater wished to do so, in effect making a nine-step scale.

In the final stage of the selection procedure, the interpretation of the letter grades was changed as follows:

A Definitely should be disseminated

B Good bet

C Will accept

D Bad bet

F Won't accept

There was general agreement that the overall rating should be a global judgment based on the criteria and subcriteria discussed under the side headings above. In addition, however, it was recognized that the rater might wish to raise or lower his rating for a particular product by taking account of USOE priorities, availability of other dissemination support for the product, his judgment of the product's potential educational impact, or his knowledge of the existence of

competitive products of equal or greater cost-effectiveness and desirability. In the work of the panel, then, these factors were treated as supplementary considerations to be taken into account after the product's characteristics had been evaluated. It was recognized, however, that there might have been advantages if these considerations had been introduced at earlier points in the evaluating process.

The foregoing formulations related to criteria grew out of discussions of a "draft for comment" prepared in advance by staff for panel consideration. This paper is included as Appendix G. The panel's initial formulation of criteria with accompanying Rating Form, used during the initial selection stage is Appendix H. Its revised criterion statement with the Rating Form to accompany it, used during the final selection stage, can be found in Appendix I.

Suggestions Related to Criteria for 1972

A careful study of the documentation on the evolution of criteria for use in product selection will show a high degree of consistency between early staff expectation and ultimate panel realization. Those involved in the process, for the most part, were well satisfied that the formulation of criteria was quite adequate for the initial cycle, and further that it would serve as an excellent point of departure for the new panel in its consideration of criterion-related questions for the 1972 selection cycle. As an indication of possible further refinements that suggest themselves after the selection, Professor Scriven has prepared a version VIII of the Rating Form, which is shown as

Figure 3. The panel suggested in its reprise that the Rating Form might make a useful tool for Project Officers at OE to use in their preliminary screening of products for inclusion in the upcoming pool.

The panel also emphasized in its reprise that the rating forms turned out to be outstandingly useful tools to help focus their review of products upon issues of evaluation. The forms proved an ideal way for panelists to collect their judgments systematically and to "store" them for later referral. Their more important strength, though, was that they facilitated the move from individual to group judgments about the product at hand.

For the one not sitting on the panel, the Formulation of Criteria paper is an essential reference, since the short-hand of the rating form may be found somewhat elusive to grasp. But the panel would have been hard pressed to apply its criteria to the process of selection in this diverse field of 51 products without the aid of its rating form. In a sense, then, the rating form may best be regarded as a paper upon which the panelist wrote a memo to himself.

As the eight rating form versions already attest, an instrument such as this should be expected to be under constant revision. In the context of its use, moreover, it should be noted that measures of interrater reliability would be inappropriate. The form, simply put, is a common framework for the making of individual judgments, which conceivably, might have very little in common. Nonetheless, an order of business, in the procedures early in 1972, might very well be a check on intra-rater reliability, using the members overlapping with the 1971 Panel as the subjects.

The only fundamental improvement related to criteria seen to be needed for 1972 is incorporation of Chairman Krathwohl's concept that the overall rating be made in the context of confidence level. A notation, for instance, that a given product rating -- whether high or low -- was made with relatively low or relatively high confidence would provide another dimension along which to seek consensus during panel discussion.

ETS-PEP-MJS 5/71

PRODUCT EVALUATION RATING FORM

PRODUCT REF. NO. . . . DESCRIPTORS . . . . . .

RATER NAME/CODE . . . . . DATE . . . . . .

GOALS Considerations (use check mark(s) to indicate any that were significant positive factors in your judgment; a cross to indicate significant lacks (or drawbacks) & a query to indicate felt unreliability of judgment)

- Urgent present need . . . . .
- Educational centrality . . . . .
- Size of target population . . . . .
- Notable originality . . . . .
- 'Leverage' or 'multiplication' effect . . . . .
- Relation to USOE priorities . . . . .

OVERALL RATING

A Excellent B Good C Fair D Poor F Undesirable ? No good estimate possible

(Circle one or more; add + if desired)

EFFECTIVENESS

Considerations

Adequacy of Test Data → Sample size . . . . . ; Sample fairness . . . . . ; Controls Adequacy . . . . . ; Validity of Criterion . . . . . ; Objectivity of Judgment . . . . . ; Follow-up . . . . . ; Background or Internal Evidence . . . . .

Actual Performance in Field Tests → Student Performance/Attitude Changes . . . . . ; Teacher Reactions . . . . .

Observed Side Effects . . . . . (on target population &/or different (untargeted) population) . . . . . Desirability of Side Effects (rate as for Goals above) A B C D F ?

OVERALL PERFORMANCE RATING ON STATED GOALS A B C D F ?

ON SIDE EFFECTS A B C D F ?

"COSTS"

Considerations

Cost of materials: basic/first-time . . . . . ; per annum . . . . . ; enrichment materials . . . . .

Extra/Less Staff: Professional . . . . ; Paraprofsnl . . . . ; Technical . . . . ; Clerical . . . .

On-Service Training Costs . . . . . ; Consultant Costs . . . . .

Repair/Upkeep/Running/Temporary Replacement Costs . . . . . ; Extra/Less Space . . . . .

Special Needs eg. projection room . . . . . ; System Disruption Costs eg. "weaning" problem . . . . . ; Opposition/Support from Students . . . . ; Staff . . . . ; Community . . . .

OVERALL COST RATING A B C D F ?

ADOPTABILITY

Considerations

Materials → Samples only . . . . . ; Limited quantities now . . . . . ; Unlimited . . . . . ; Tests &c also available . . . . ; Enrichment materials . . . . . ; Situation unclear . . . . .

Use → Requires advance training . . . . . ; Requires plant changes . . . . . ; Requires support system contracts . . . . .

OVERALL ADOPTABILITY RATING A B C D F ?

FINAL Rating on Presently Demonstrated Desirability of USOE Dissemination Support

Considerations

- Preceding Judgments of → Goals . . . . . ; Effectiveness . . . . ; Costs . . . . . ; Adoptability . . . . .
- Availability/Unavailability of Other Support . . . . . .
- Availability of Critically Competitive Products . . . . . (Identify: . . . . . .)

A B C D F ? Hold for next year's pool.

Further comments; . . . .

Figure 3 - Rating Form VIII

PROCEDURES IN INITIAL AND FINAL STAGES OF SELECTION

Introduction In devising the procedures for accomplishing its task, the panel established a general strategy for product evaluation and then made detailed decisions when the need for them arose. There appeared to be a strong consensus on two main points of procedure:

(1) Before making a judgment regarding a product, each panel member considered it necessary to familiarize himself with both the precis and other supporting information concerning the product, consulting the dossier of source material when necessary. Procedures calling for a more analytical approach to the evaluation task were judged to be unsuitable.

(2) Each panel member had full responsibility for achieving an overall judgment with respect to each product that he rated, using the available information in whatever way he thought best.

The major staff assistance to the panel in setting procedures for the two selection stages took the form of a working paper sent panel members for study and advance consideration prior to the first selection meeting. The plan suggested a general strategy for making panel decisions and a set of alternatives for panel consideration in the course of reaching them. Procedures suggested by the staff will be found in Appendix J.

Initial Stage of Selection The general plan provided for an initial selection stage in which to identify a limited number of products for intensive study. This initial selection was accomplished at the second of the three two-day panel meetings, utilizing a basic plan developed at the initial meeting and the rating criteria shown as Appendix H. The panel consensus was that if each panel member read and evaluated approximately half of the products, there would be sufficient time remaining to arrive at some form of agreement on an overall rating for each product and to establish standards for selecting a limited number of products to be considered at the final selection stage.

In the initial selection stage, the panel was divided into two working subpanels. An attempt was made to balance the two subpanels with respect to the kinds of educational experience and expertise represented. (One member of each subpanel was absent, but a substitute member was provided.) Odd-numbered products were assigned to one subpanel for evaluation; even-numbered products to the other. Each subpanel had its own workroom for reading and discussion. Dossiers for its products were available for consultation in its room. Plans were devised to permit rotation of subpanel members, but the subpanels decided not to use them.

The activities involved in accomplishing the initial selection included four main steps, as follows:

(1) As soon as all members of a subpanel had read and rated the first six products, the subpanel discussed the ratings given to each product by each rater and arrived at a

consensus. (In one subpanel, each product was discussed by subpanel members and a subpanel rating was agreed upon. In the other subpanel, each member rated each product again after discussion and the sum of the revised ratings, no matter how disparate, was taken as the consensus.)

Each subpanel agreed that it would discuss the first six products after rating them, to permit subpanel members to exchange ideas on the basis of specific products and to obtain a preliminary indication of how difficult it would be to achieve consensus. It was at this time also that each subpanel reached a decision not to adopt any plan for subpanel rotation, but to stay together as a working group.

(2) Each subpanel then evaluated the remainder of its products, reached consensus in the same way as it did for the first six, and made a tentative decision about whether or not each of its products should be included in the final stage of selection.

(3) Each subpanel rated three additional products selected arbitrarily from the list of products for which the other subpanel was responsible. This step made it possible for the entire panel to discuss the six products in common and to clarify differences in viewpoint. As it turned out, most of the discussion centered on two of the six products. Each subpanel had selected one product for retention which the other subpanel had agreed to screen out.

(4) On the basis of the discussion of the products rated by both subpanels, and a consideration of differences in the

number of products recommended by each subpanel for retention, it was decided that each subpanel should be responsible for deciding whether the products initially assigned to it should be retained for consideration in the final stage of selection. This decision was considered "fail safe" in that the selection decision at the final stage would be made by the entire panel. It should be noted that the procedure followed does not preclude the possibility that the panel might have recommended a product that had been removed by a subpanel from final consideration. Provision was made for recommendation by a subpanel that the initial selections be made by the panel. Neither subpanel found it worthwhile to exercise this option. It appears that each subpanel recommended one or more products for the final stage because it was judged that certain ambiguities in the available information might, on further study, be resolved in a way that would justify recommendation of the product for dissemination.

On the whole, then, a product passed the initial screening if the subpanel to which it had been assigned judged it to have a reasonable possibility of being recommended for dissemination at the final stage of selection.

Final Selection The final selection procedures were designed with relatively little discussion, in part because the Panel had reached a consensus on the main procedural issues and in part because the

final selection decisions were made by all seven members present at the third session of the panel. It was agreed that an attempt to reach consensus on the weighting of separate criteria would not be desirable.* Instead, evaluative comments made during the discussion of a product or written on the rating sheets (see Appendix I) were thought to provide a better way of communicating the panel's evaluation of a product to USOE. A further decision redesignated the five rating levels as follows:

A Definitely should be disseminated

B Good bet

C Will accept ("will live with it")

D Bad bet

F Won't accept ("will fight it")

The procedures for the final selection stage were as follows:

(1) Each panel member, using the criteria and a rating sheet as a guide, worked through the materials in his book (precis, reviewer's comments, site-visit report) on each of the 20 products, and dossiers were available as needed. The purpose of this step was to assign an overall rating to each product.

(2) The full panel then met to discuss and re-rate all 20

* Four panelists did, as a presession independent exercise, apply weights to each of the four criteria, to the considerations within each, and to the incremental additions in the overall rating. This was done by distribution of 100 percentage points. Agreement by the four was within 10 points on "Goals" and "Effectiveness," 15 points on "Costs," and 5 points on the Overall residuals. The four panelists who abstained from this exercise regarded such weighting to be misleading, and the concept was dropped.

products. Typically, the chairman polled the members for initial ratings. Members who had assigned a relatively high rating made their evaluative comments first, followed by members who had given a low rating. In most instances, panel members then rated the product again, making out a revised rating sheet if necessary. A new polling determined the semifinal ratings.

(3) It was agreed that the revised panel ratings could be summed to yield a total score for each product, using a 5-point scale with A=5, B=4, and so forth. Abstentions were handled by substituting the average of the ratings for the missing rating. (In two instances, a tentative rating of "Hold" had been assigned. The rater, however, replaced these ratings with ratings on the usual scale.)

(4) Those products that had received relatively high total ratings were listed in rank order on a blackboard, a tentative cutting score was agreed upon, and there was further discussion of products in the vicinity of this score and some adjustments in ratings of products near the threshold.

(5) At this point, a final set of ratings was made by the panel, the results were tabulated, and a firm decision on the cutting score was made. It turned out a product needed a total rating of 20 (an average rating of 2.9) or higher in order to be selected at the final stage. Thus, a product rated "3" (will accept) by 6 raters and "2" (bad bet) by

one rater would have qualified for selection. It was the hope, however, that the cutting score would have the effect of eliminating a product that lacked a vote of "4" (good bet). Had a cutting score of 20 allowed such a product in, the threshold likely would have been moved upward.

The use of a rating scale formulated in terms of the action to be taken by the panel and the use of a numerical method for combining ratings for different raters worked very well and appears to have facilitated the setting of a standard for acceptance.

Panel decisions on final selection procedures went so rapidly in part because an essay on the subject prepared at staff request by Professor Stake (who was not present) had been distributed in advance and a companion piece by Professor Scriven had been handed out at the start of the session. These papers are included in Appendix K.

Refinements in Selection Procedures for 1972 The major change that should be made in selection procedures would effect the initial selection phase. Two days is an insufficient period of time for reviewing as many as 50 products on a systematic and methodical basis. As a practical matter, the 1971 panel had no alternative but to work as two subpanels without member rotation. There was not that much time margin to allow moving panelists from one subpanel to another on a preplanned basis. The change suggested would schedule the initial selection session (panel meeting II) for three days, and implement one of the plans for panel-member rotation suggested in Appendix J (pp. 2-4). The major result of this change would be to eliminate significant differences on

a stringency-leniency scale that otherwise would occur were panelists not rotated.

A second change -- this at the final selection stage -- would involve the order in which the information on initially selected products appears in each panelist's final-stage book. Thus the description of product X would be read first by one panelist, third by another, last by another, and so on. This change would keep the panelists' judgment free of any bias caused by timing.

As noted elsewhere, a book-order numbering system -- especially at the initial stage -- would add considerable flexibility to the panel's use of the available information. If the shuffling scheme described above is not used for the final stage, a book-order numbering system and an "order of reading" assignment sheet for each panel member will do as well to break up consistent patterns of reading order.

There was one problem the panel recognized, but did not have time to solve. The panelists were concerned that their decision about the fate of a product might be construed as simply a yes-or-no matter, when actually, there must be at least several levels of action a panel with this function should take regarding a product. Along one dimension, levels of action might follow a hierarchy of conditions from "simple to satisfy" to "difficult to satisfy." For example, Product X could be disseminated now if the validation study due next month reports significant gains. Or, Product X might be disseminated next year if a summative evaluation shows change attributable to the product.

In planning for improvements in procedure, the panel members urged that levels of recommendation be suggested within which their judgments regarding a given product and its future development, field testing, and use could be applied. The effect of this change should be to make it possible to recommend limited dissemination -- installation in a few places, acquisition of independently collected data -- before massive dissemination is contemplated.

There was some interest in the development of a "product profile," an instrument that would (a) force principal investigators reporting on their products to communicate in criterion-related terms, and (b) structure evaluative feedback to principal investigators that would reflect criterion-related refinements. The product profile would communicate both on what had been done and what, in the view of disinterested parties, should next be done. As in the case immediately above, time was unavailable for fuller development of this idea.

SUMMARY OF 1971 INITIAL AND FINAL SELECTIONS

Introduction From a field of 51 products in the 1971 pool, 20 were initially selected for analysis-in-depth, and nine of those are recommended for dissemination actions by NCEC. The list of selections at each stage is reasonably representative of the whole field both insofar as major purpose is concerned and as to target, field, and level. There are concentrations in both lists of materials developed for the improvement of practices among teachers in service.

Twenty Products Screened for Further Review The initial selection meeting of the 1971 product evaluation panel led to more detailed review and analysis, the expectation being that from this group, the panel would make final selections of those products to be the subjects of NCEC's focused dissemination program.

Products are listed in "book order," according to the sequence set by the system of classification. Thus, among the twenty are two products whose purposes deal with ending neglect, two classed as individualized instruction, four as group instruction, seven for upgrading the quality of teaching, two related to administration, and three whose focus is on at least two of the major program areas noted above.

Nine of the products come from regional educational

laboratories, four from research and development centers and seven from colleges, universities and other independent investigators.

The 20 products initially selected, together with the sites of the developments and the principal investigators are shown below.

AC 58 Tab 1 Classification 1219 Site: Austin, Texas
Bilingual Early Childhood Education Learning System
SWEDL (Nedler)

AC 45 Tab 1 Classification 1419 Site: Albuquerque, N.M.
Reinforced Readiness Requisites Program
SWCEL (Olivero - Speiss)

AC 01 Tab 2 Classification 2X37 Site: Chicago, Illinois
Job Experience Kits
Stanford University (Krumboltz) Now SRA

AC 25 Tab 2 Classification 2X59 Site: Boston, Mass.
MATCH Box
Boston Children's Museum (Kresse)

AC 67 Tab 3 Classification 3X11R Site: Inglewood, Calif.
Communications Skills Program
SWLERD (O'Hara)

AC 11 Tab 3 Classification 3X35 Site: Minneapolis, Minn.
Social Studies Curriculum Guides and Materials K-12
University of Minnesota (West)

AC 15 Tab 3 Classification 3X55 Site: Providence, R. I.
Geo Historical Structure for Social Studies Curriculum
Rhode Island College (Shinn)

AC 52 Tab 3 Classification 3X37 Site: Columbus, Ohio
Industrial Arts Curriculum Project
Ohio State University (Lux and Ray)

AC 33 Tab 4 Classification 4449 Site: Kansas City, Mo.
Cooperative Urban Teacher Education (CUTE)
McREL (Clothier)

AC 68 Tab 4 Classification 4940 Site: Berkeley (U.) Calif.
Instruments and Procedures for Describing Effective Teacher Behavior
U. of California (Berkeley) (Medsker)

AC 64 Tab 4 Classification 4X43 Site: Austin (U.), Texas
Teaching of Science: A Self-Directed Personalized Teacher Education Program
U. of Texas (Butts and Hall)

AC 28 Tab 4 Classification 4X44R Site: Berkeley, Calif.
Individualizing Instruction in Mathematics (Minicourse 5)
FWLERD (Gall)

AC 19 Tab 4 Classification 4X40 Site: Columbus, Ohio
Simulation Training in Planning Vocational Education Programs and Facilities
Ohio State U. (Ward)

AC 38 Tab 4 Classification 4X40 Site: Minneapolis, Minn.
Individualized Instruction Through Contingency Management
UMREL (Morreau)

AC 35 Tab 4 Classification 4X40 Site: Portland, Oregon
Instructional System in Development of Higher Level Thinking Abilities
NWREL (Fish)

AC 22 Tab 5 Classification 541X Site: Maryville, Mo.
Individual Readiness Test
Northwest Missouri State College (Walker)

AC 43 Tab 5 Classification 5519 Site: Berkeley, Calif.
Parent/Child Toy Lending Library
FWLERD (Nimnicht, Brown, Johnson, Addison)

AC 24 Tab 6 Classification 6519 Site: New York City
Educational Television for Preschoolers (Sesame Street)
Children's Television Workshop (Ganz)

AC 70 Tab 6 Classification 6550R Site: Madison, Wisconsin
Multi-Unit Elementary School
WREDC/CL (Klausmeier)

AC 42 Tab 6 Classification 6940R Site: Berkeley, Calif.
Effective Questioning-Elementary Level (Minicourse 1)
FWLERD (Borg)

Nine Products Recommended for Dissemination There follows a list of nine products recommended to make up the 1971 validated product pool. These products are recommended by the Product Evaluation Panel for special focused dissemination attention by the National Center for Educational Communication.

The products are listed in descending order from the highest rating received. As is described in detail elsewhere, each panelist, after applying the criteria for product evaluation, gave the product an overall rating on a scale

A - B - C - D - F

5 - 4 - 3 - 2 - 1

These ratings for seven panelists were summed and the total served to yield an initial ranking. Panel plenary reviews led in some cases to adjustments in individual ratings and panel totals. The sum of ratings on the nine products in the recommended list clustered between 29 (B+) and 20 (C-).

No serious reservations were registered with respect to AC42, AC43, AC70, or AC33 though there were a number of constructive suggestions regarding their improvement in the course of dissemination, and these are reported elsewhere.

A review of definitive data to become available during summer of 1971 should precede explicit plans on the dissemination of AC64. Data were not there at panel selection time.

With regard to AC25, the major need for "upgrading" before dissemination is for (a) clearer definition of product goals and (b) closer alignment of product components to the goals, so that the potential user may readily perceive what the components do to help advance the objectives.

"The House of Greece" was pointed to as a model for other components to emulate.

AC28 was thought to be in need of better evaluative data, ideally including observed changes in students or in teacher behavior, hopefully with controls, minimally with comparisons. The same general comment applies to AC45: data are presented only on one of the stated objectives and data on the others are needed; there are as yet no findings that the product relates to later reading achievement, the function of "readiness" needs to be established.

In the case of AC67, the absence of summative evaluation data soon to become available caused the panel to place the product in a marginal category, at one point in a "Hold for 1972 verification." That the product seemed to have definite promise, however, caused the panel to place the product on the threshold rather than below it.

First Priority Recommendations - (All Panel Ratings were either A, B or C.)

AC 42 Tab 6 Classification 6940R Site: Berkeley, Calif.
Effective Questioning-Elementary Level (Minicourse 1)
FWLERD (Borg)

AC 43 Tab 5 Classification 5519 Site: Berkeley, Calif.
Parent/Child Toy Lending Library
FWLERD (Nimnicht, Brown, Johnson, Addison)

AC 70 Tab 6 Classification 6550R Site: Madison, Wis.
Multi-Unit Elementary School
WREDC/CL (Klausmeier)

AC 33 Tab 4 Classification 4449 Site: Kansas City, Mo.
Cooperative Urban Teacher Education (CUTE)
McREL (Clothier)

AC 64 Tab 4 Classification 4X43 Site: Austin (Univ), Texas
Teaching of Science: A Self-Directed Personalized Teacher Education Program
U. of Texas (Butts and Hall)

Second Priority Recommendations - (At least one Panel Rating of D or lower.)

AC 25 Tab 2 Classification 2X59 Site: Boston, Mass.
Match Box
Boston Children's Museum (Kresse)

AC 28 Tab 4 Classification 4X44R Site: Berkeley, Calif.
Individualizing Instruction in Mathematics (Minicourse 5)
FWLERD (Gall)

AC 45 Tab 1 Classification 1419 Site: Albuquerque, N.M.
Reinforced Readiness Requisites Program
SWCEL (Olivero - Speiss)

AC 67 Tab 3 Classification 3X11R Site: Inglewood, Calif.
Communications Skills Program
SWLERD (O'Hara)

RECOMMENDATIONS FOR 1972

Introduction In the last part of each major section of this report attempts are made to evaluate the 1971 activities and to suggest needed improvements for 1972. Taken together, then, these concluding remarks, scattered throughout the report according to the topic to which they pertain, constitute our recommendations for 1972. Although some of the suggestions have been made by staff members of either EPIE or ETS, most have been subjected to scrutiny by the panel, and a large proportion of the suggestions have come from panel members.

In this section, highlights of the recommendations are repeated. Details will be found in the section covering the subject to which the recommendation is related.

On Products for the Pool

- Products should be entered into the pool as they become ready; the pool should be on-going
- Developers should describe product verifications against a backdrop of product goals and objectives
- A comprehensive analysis of likely user-costs should be an integral part of the product description - both direct and indirect costs
- Selection criteria for this project should be used as guidelines to increase the usefulness of product descriptions
- Invitations for entries into the pool should

foster nomination of promising products, deter nomination of products inappropriate in the context of this mission

- o An earlier start for the process is needed
- o Product nominations also should be sought outside the present sphere of OE-sponsored research

On Establishing the Panel

- o The 1971 panel composition sets a good standard for 1972
- o Organization of the panel should be such as to encourage much of its work to be done in plenary sessions
- o The membership of the panel should be enlarged to ten, adding a classroom teacher and a school district research director
- o There should be an overlap of two or three panel members in successive years

On Initial Information on Products

- o More elapsed time should be scheduled for initial study of product-related documentation and preparation of product-precis
- o An especially trained precis writing team should be organized to become highly knowledgeable both in this type of "journalism" and the substantive elements of product development and verification.
- o Selection criteria should serve as the key guide in the preparation of product precis

- o Quality control procedures for precis and other product-related descriptions should be instituted
- o Product identification should be made easier, quicker, and more functional

On Additions to Product-Related Information

- o Site visits if at all possible should be made by precis writers; they should use a structured interview guide
- o A final expanded precis should be done to subsume all information useful to the panel on a given product in a single document
- o Product reviewers and analysts should be readily available to appear if needed as "expert witnesses" when the panel selections are being made

On Selection Criteria

- o Further improvement should take fullest possible advantage of the evolutionary development of the 1971 selection criteria and the instruments for applying them to product evaluation
- o The product evaluation rating form shown as Figure 3 should serve as the point of departure for considering criterion-related questions in 1972
- o The selection criteria should be considered for their possible usefulness in the preliminary screening of products to be entered into the pool and in preparing product descriptions for panel consumption

- o A rating form should continue to be the framework within which criteria are applied to product evaluation; the form to be used in 1972, however, should be subjected to an intra-rater reliability check prior to being put into operational use

On Selection Procedures

- o The initial selection meeting of the panel should be scheduled for three days; panel members should be rotated so that each panel member is enabled to review and evaluate products with every other panel member
- o At the final selection stage products to be reviewed and evaluated should be shuffled so that each panel member conducts his evaluations in a sequence different from the others
- o Selections should not be expected only on a yes-no basis; it should be possible to conditionally select, and for the conditions to be graduated on a scale from "much to be done prior to dissemination" to only a "modest amount of further work called for prior to dissemination
- o An instrument to guide the development of "product profiles" should be devised as a base for two-way communication to convey useful information about a product; it would aid developers in supplying needed information on what has been done to users and evaluators,

it would aid the latter group in supplying information to the former on what still needs doing

- As an augmentation of an overall rating on a product by a panel member, a confidence level scale should be added; the product-rating would then fit into a context varying in accordance with the adequacy of the information supplied and the capacity of the evaluator to assimilate it

APPENDICES

February 1, 1971

Appendix A

PEP Accession List

01 Job Experience Kits 7-0111 2X37
Dr. John Krumboltz, Principal Investigator
Stanford University

02 The Vocational Development Inventory 5-0038 5X57
Dr. John O. Crites, Principal Investigator
University of Iowa

03 Project TALENT Data Bank 5-0606 5XXX
Dr. John G. Claudy, Principal Investigator
American Institute for Research

04 Harvard Project Physics 5-1038 3X33
Professor Gerard Holton, Associate Professor
F. James Rutherford, Professor Fletcher G. Watson,
Principal Investigators
Harvard University

05 A System for Individualizing and Optimizing Learning Through Computer Management of the Educational Process
8-0157 5X5X
Dr. Alexander Schure, Principal Investigator
New York Institute of Technology

06 Illinois State-Wide Curriculum Study Center in the Preparation of Secondary School English Teachers
HE-145 5X41
Professor J. N. Hook, Principal Investigator
University of Illinois

07 English Open to All Junior and Senior High School Students
HE-080 3X31
Edward B. Jenkenson, Principal Investigator
Indiana University

08 Inquiry Materials for Social Studies HS-041, H-292 3535
Edwin Fenton, Principal Investigator
Carnegie-Mellon University

09 Multi-Media Economics Curriculum Development Project
8-0447 2X45
Edmund W. Fitzpatrick, Principal Investigator
Educational Technology Center

10 Intermediate Science Curriculum Project (ISCS) 6-1762 2X33
Ernest Burkman, Principal Investigator
Florida State University

11 Preparation and Evaluation of Social Studies Curriculum Guides and Material for Grades K-14 5-0659 3X35
Edith West, Principal Investigator
University of Minnesota

12 The Taba Curriculum Development Project in Social Studies
5-1314 3X55
Dr. Norman Wallen, Principal Investigator
San Francisco State College

13 Preparation for a Dual Role: Homemaker - Wage Earner 3837
7-0006 (Part A)
Dr. Phyllis Lowe, Principal Investigator
Purdue University
6-3050 (Part B)
Dr. Helen Nelson, Principal Investigator
Cornell University
6-3049 (Part C)
Dr. Julia Dalrymple, Principal Investigator
Ohio State University

14 New Careers in Public Service 7-0192 5X57
Randy H. Hamilton, Principal Investigator
Institute for Local Self-Government

15 A Study of Geo-Historical Structure for a Social Studies Curriculum 6-1195 3X55
Ridgway Shinn, Jr., Principal Investigator
Rhode Island College

16 Project Africa 7-0724 1139
Barry K. Beyer, Principal Investigator
Carnegie-Mellon University

17 The Development of Instructional Materials and Teaching Strategies on Race and Culture in American Life 8-0197 3X25
Dr. John S. Gibson, Principal Investigator
Tufts University

18 A Regional Study of the Aviation Mechanics Occupation 5-0189
3837
David Allen, Principal Investigator
University of California (Los Angeles)

19 Simulation Training in Planning Vocational Educational Programs and Facilities 7-0158 4X40
Dr. Darrell L. Ward, Principal Investigator
Ohio State University

20 The Oregon Curriculum: A Sequential Program in English 5-0366 3X31
Albert R. Kitzhaber, Principal Investigator
University of Oregon

21 Drug Education Program 9-G-067 3X59
Vidal M. Trevino, Principal Investigator
Laredo Independent School District, Laredo, Texas

22 Individual Readiness Test 9-F-017 541X
Wanda Walker, Principal Investigator
Northwest Missouri State College

23 Unified Mathematics Program 7-0711 3634
Dr. Howard F. Fehr
Columbia Teachers College

24 Educational Television for Preschoolers (Sesame Street) 8-0475 6516
Mrs. Joan Ganz, Principal Investigator
Children's Television Workshop

25 MATCH Box 5-0710 2X59
Mr. Frederick H. Kresse, Principal Investigator
The Children's Museum

26 Behavioral Objectives Package 4740
Dr. James L. Olivero, Dr. Carmen R. Rimiraos, Principal Investigators
Southwestern Cooperative Educational Laboratory

27 Backgrounds in Language 4X41
Mrs. Barbara K. Long, Principal Investigator
Upper Midwest Regional Educational Laboratory

28 Individualizing Instruction in Mathematics (minicourse 5) 4X44
Dr. Meredith D. Gall, Principal Investigator
Far West Laboratory for Educational Research and Development

29 Instructional System in Research Utilizing Problem Solving 4X40
Lawrence D. Fish, Principal Investigator
Northwest Regional Educational Laboratory

30 Improving Motor-Perceptual Skills 4X50
Lawrence D. Fish, Principal Investigator
Northwest Regional Educational Laboratory

31 Instructional System in Interpersonal Communications 4X40
Lawrence D. Fish, Principal Investigator
Northwest Regional Educational Laboratory

32 Coordinated Helps in Language Development 4X40
Lawrence D. Fish, Principal Investigator
Northwest Regional Educational Laboratory

33 Cooperative Urban Teacher Education 4449
Dr. Grant Clothier, Principal Investigator
Mid-Continent Regional Educational Laboratory

34 Instructional System in Classroom Questionning Strategies 4X40
Lawrence D. Fish, Principal Investigator
Northwest Regional Educational Laboratory

35 Instructional System in Development of Higher Level Thinking Abilities 4X40
Lawrence D. Fish, Principal Investigator
Northwest Regional Educational Laboratory

36 Dialects and Dialect Learning - An English Inservice Program 4X41
Karen Matison Hess, Principal Investigator
Upper Midwest Regional Educational Laboratory

37 Learning Standard English 4X40
Charles A. Findley, Karen Matison Hess, Principal Investigators
Upper Midwest Regional Educational Laboratory

38 Individualized Instruction Through Contingency Management 4X40
Mr. Lanny E. Morreau, Principal Investigator
Upper Midwest Regional Educational Laboratory

39 Instructional System In Facilitating Inquiry in the Classroom 4X40
Lawrence D. Fish, Principal Investigator
Northwest Regional Educational Laboratory

40 The Inquiry Role Approach Component of the Development of Inquiry Skills Program 2X30
Richard M. Bingman, Principal Investigator
Mid-Continent Regional Educational Laboratory

41 Model CM Classrooms for Individualized Instruction - Grades 4-6 2X20
John C. Maxwell, Principal Investigator
Upper Midwest Educational Laboratory

42 Effective Questionning - Elementary Level (minicourse 1) 4X40
Dr. Walter R. Borg, Principal Investigator
Far West Laboratory for Educational Research and Developme nt

43 Parent/Child Toy Lending Library 5519
Glen Nimnicht, Edna Brown, Stan Johnson, Bertha Addison Principal Investigators
Far West Laboratory for Educational Research and Development

44 The Oral Language Program 1211
Dr. James L. Olivero, Dr. Robert T. Reeback, Mrs. Helgi Osterriech, Principal Investigators
Southwestern Cooperative Educational Laboratory

45 Reinforced Readiness Requisites Program 1419
Dr. James L. Olivero, Mrs. Madeleine Speiss, Principal Investigators
Southwestern Cooperative Educational Laboratory

46 English as a Second Language 4741
Dr. James L. Olivero, Dr. Carmen R. Timiraos, Principal Investigators
Southwestern Cooperative Educational Laboratory

47 Self-Instructional System in Basic Electricity 2X37
Lawrence D. Fish, Principal Investigator
Northwest Regional Educational Laboratory

48 Instructional System in Systematic and Objective Analysis of Instruction 4X40
Lawrence D. Fish, Principal Investigator
Northwest Regional Educational Laboratory

49 Instructional System in Interaction Analysis 4X40
Lawrence D. Fish, Principal Investigator
Northwest Regional Educatonal Laboratory

50 Self-Instructional System in Welding 2X37
Lawrence D. Fish, Principal Investigator
Northwest Regional Educational Laboratory

51 Instructional System in Speech 6X31
Lawrence D. Fish, Principal Investigator
Northwest Regional Educational Laboratory

52 Industrial Arts Curriculum Project 7-0003 3X37
Donald G. Lux, Willis Ray, Principal Investigators
Ohio State University

53 An Enlarged Music Repertory for Kindergarten Through Grade Six 5-0219 3X58
Gordon Hardy, Principal Investigator
The Julliard School

54 Academic Building Systems 809113 5X40
R. Clayton Kantz, Principal Investigator
University of California (Berkeley)

55 Social Service Aide Project (Career Options Research & Development) 7-0329 1437
Robert K. Soong, Principal Investigator, parts 1 and 2
Miss Jean Wetzel, Principal Investiagtor, parts 3, 4, 9, 10, 11
Barry S. Warren, Principal Investigator, parts 5, 6, 13, 14
George A. Kich, Principal Investigator, part 8
YMCA of Metropolitan Chicago

56 Western Interstate Commission for Higher Education, Planning, Analysis and Management Systems Project 8-0708 3940
Dr. Robert Huff, Principal Investigator
WICHE/MPS

57 A Comprehensive Curriculum in Dance for Secondary Schools 5-0244 3X38
Nadia Chilkovsky Nahumck
Philadelphia Dance Academy

58 Bilingual Early Childhood Education Learning System 1219
Mrs. Shari Nedler
Southwest Educational Development Laboratory

59 Multicultural Social Education Program 1415
Mrs. Martha Smith
Southwest Educational Development Laboratory

60 Research and Development on Preschool Disadvantaged Children 5-1181 1410
Professor Merle B. Karnes
Institute for Research on Exceptional Children

61 Exploration in Biology Topics (Inquiry Skills Program) 5X33
Eugenia M. Koos
Mid-Continent Regional Educational Laboratory

62 Development of Materials for a One-Year Course in African Music for the General Undergraduate Student 6-1179 1148
Dr. Vada E. Butcher, Principal Investigator
Howard University

63 A Program for Leadership Training in Team Teaching 4440
L. Jean York, Principal Investigator
University of Texas (Austin)

64 The Teaching of Science: A Self-Directed Personalized Teacher Educatton Program 4X43
David Butts and Gene Hall, Co-Investigators
University of Texas (Austin)

65 Comprehensive Personal Assessment and Counseling Feedback Systems for Pre-Service Teacher Education Programs 4940
Oliver H. Brown, Co-Director
University of Texas (Austin)

66 Alternatives for Learning Through Educational Research and Technology (ALERT): An Educational Information System 5959
C. L. Hutchins, Principal Investigator
Far West Laboratory for Educational Research and Development

67 Communication Skills Program 3X11
Robert W. O'Hare
Southwest Regional Laboratory

68 Instruments and Procedures for Describing Effective Teacher Behavior 4940
Leland L. Medsker, Director
University of California (Berkeley)

69 An Instrument and Procedures for Improving Communication and Academic Policy Making 5940
Leland L. Medsker, Director
University of California (Berkeley)

70 **Multi-Unit Elementary School**

Dr. Herbert J. Klausmeier, Principal Investigator
Wisconsin R & D Center for Cognitive Learning

Appendix B

Project on the Evaluation of Products

Products Suggested for Deferral to 1972

AC03 5XXX
Project TALENT Data Bank BR 5-0606
American Institutes for Research

Only remotely related to explicit improvements in educational practices.

AC07 3X31
English Open to All Junior and Senior High School Students HE 080
Indiana University

Report due out April 15, 1971. Access to final report and its summary of evaluative comment is essential. Hard verification data presently lacking.

AC13 3837
Preparation for a Dual Role: Homemaker-Wage Earner BR 6-3049, BR 6-3050, BR 7-0006
Cornell, Ohio State, Purdue Universities

Report due Spring 1971. Results on effectiveness needed.

AC14 5X57
New Careers in the Public Service: A Model for Redirection of Vocational-Technical Education BR 7-0192
Institute for Local Self-Government

Final Report due June 1971. Essential for analysis of verification data.

AC18 3837
The Aviation Mechanics Occupation BR 5-0189
University of California at Los Angeles

Inappropriate to include. The project has too specific a target to be of interest to NCEC in the context of dissemination attention (17,000 students, 130 schools makes up the target).

AC30 4X50
Improving Motor-Perceptual Skills
Northwest Regional Educational Laboratory

Longitudinal evaluation is being done, no due date given. Results needed.

AC37 4X40
Learning Standard English
Upper Midwest Regional Educational Laboratory

Prototype system in only two classrooms. Field test not yet available. Final report August 1971.

AC54 5X40
Academic Building Systems BR 8-0113
University of California (Berkeley)

Will not be feasible to do analysis of this product without full report which is due April 1971. The product may even then fall outside the purview of the current effort since its emphasis is remote from the instructional process.

AC55 1437
New Careers Research, Social Service Aide Project BR 7-0329
YMCA of Metropolitan Chicago

To maintain its integrity, this whole project, containing some fourteen sub-products should be looked at at one time, probably in 1972. Some of the sub-tasks are barely off the ground. Reporting dates range from January 1971 (sub-products) to October 1971, with nine of the fourteen due not until Fall 1971.

Wesley W. Walton, Ed D
Project Director

January 29, 1971

Educational Testing Service
Princeton, New Jersey

Project on the Evaluation of Products

Products Suggested for Deferral to 1972

DEFERRAL LIST NUMBER 2

AC06 5X41
Illinois State-Wide Curriculum Study Center in the Preparation of Secondary School English Teachers (ISCPET)
University of Illinois

The product seems far enough from high priority concerns in education and from generalizability to other states in any case as to be unlikely to lead to benefits from special NCEC attention.

AC21 3X59
Drug Education Program
Laredo Independent School District

There are eight sample drug education curriculums from throughout the country now available to guide teachers in developing local programs, and a national clearing house in Chevy Chase, Maryland to service requests. This product would appear to duplicate already available materials. There are no verification data available on the basis of which to distinguish this particular program (540 pages!) over others. Question is raised on appropriateness of this entry in terms of initial announced criteria.

AC29 4X40
Instructional System in Research Utilizing Problem Solving
Northwest Regional Educational Laboratory

Appears to be a cinch for 1972, extremely difficult to deal with before April 1971 when Technical Report becomes available.

AC31 4X40
Instructional System in Interpersonal Communications
Northwest Regional Educational Laboratory

Benefits seem well established in broad terms, will be interesting to see verification information when they get down to cases, especially in re: observable changes in interpersonal communications. Deferral pending availability of such data.

AC32 4X40
Coordinated Helps in Language Development
Northwest Regional Educational Laboratory

Verification data not readily available

AC40 2X30
The Inquiry Role Approach Component of the Development of Inquiry Skills Program
Mid-Continent Regional Educational Laboratory

This product has thus far gone through an initial feasibility study and a pilot phase during which modifications of the basic concept were made. A formative evaluation is in progress in the current year. Summative evaluation, which should precede dissemination action, will occur during 1972-73. It would seem to be a disservice to the concept to deal now with those partial products from the overall effort that are in fact now ready to be disseminated. The product appears to have been entered prematurely.

AC46 4741
English as a Second Language
Southwest Cooperative Educational Laboratory

To tell the verification story about this product would seem to require data on teachers' uses of it with students and its impact on students. Perhaps that is what their Longitudinal Study will do, but this is not presently available. Deferral pending availability of such data.

AC51 6X31
Self Instructional System in Speech
Northwest Regional Educational Laboratory

Verification data sparse, sample size too small and exposure too limited during field test. No indication that more field testing is planned. Cost of installation appears prohibitively high for target school districts. Question is raised on appropriateness of this entry in terms of the initial announced criteria.

AC57 3X38
A Comprehensive Curriculum in Dance
Philadelphia Dance Academy

Limited arena for development and testing, verification design inadequate. Inclusion of this product for later consideration should have a prerequisite of carefully designed and executed verification procedure against clear cut statements of learning objectives. Question is raised on appropriateness of this entry in terms of the initial announced criteria.

Wesley W. Walton, Ed D
Project Director

February 9, 1971

Educational Testing Service
Princeton, New Jersey

Appendix C

PRODUCT EVALUATION PROJECT

PEP

FIFTY-ONE PRODUCTS INCLUDED IN THE 1971 PRODUCT POOL FOR COMPARATIVE EVALUATION

The Product Pool from which Products will be Selected and Recommended to NCEC for Dissemination Attention

A task for the National Center for Educational Communication (NCEC), United States Office of Education, through ERIC-TM.

Educational Testing Service
Princeton, New Jersey
March 3, 1971

There follows a list of the fifty-one (51) products included in the 1971 Product Pool for Comparative Evaluation. This is the initial pool from which products will be selected for special dissemination attention by the National Center for Educational Communication.

1971 Product Pool

Product Classification

Products entered in the 1971 Product Pool for comparative evaluation have been categorized in ways that will allow comparison of the material. Each product has been classified along dimensions as noted in the four columns below. A four-digit code, then, describes for a given product its major thrust, target population, educational level, and academic field.

1971 Pool	Product Mostly for	Target Population	Educational Level	Academic Field
7	1 Ending Neglect	X1 Blacks	XX1 Pre-School & Early School	XXX1 Language Arts
7	2 Individualized Instruction	X2 Mexican-Latin Americans, Indians, Eskimos	XX2 Middle School	XXX2 Foreign Language
12	3 Group Instruction	X3 Disadvantaged Whites	XX3 Junior & Senior High School	XXX3 Science
15	4 Teacher Training	X4 All Disadvantaged Groups	XX4 College	XXX4 Mathematics
7	5 Administering Programs	X5 Disadvantaged and Other Learners	XX5 Cross Level	XXX5 Social Studies
3	6 Two or More Programs	X6 Academically Talented	XX6 No Level	XXX6 Humanities
	7 Something Else	X7 Adults		XXX7 Occupational/ Technical
	0 Undetermined	X8 Students in a Field		XXX8 Fine & Performing Arts
51		X9 Common Interest		XXX9 Multi-Disciplinary Cross-Field
		XX No Target		XXX0 No Field (A method, technique, system, or classroom management)

1. *Products Mostly for Ending Neglect

AC16 *1139
Project Africa 7-0724
Barry K. Beyer, Principal Investigator
Carnegie-Mellon University

AC44 1211
The Oral Language Program
Dr. James L. Olivero, Dr. Robert T. Reeback,
Mrs. Helgi Osterriech, Principal Investigators
Southwestern Cooperative Educational Laboratory

AC45 1419
Reinforced Readiness Requisites Program
Dr. James L. Olivero, Mrs. Madeleine Speiss,
Principal Investigators
Southwestern Cooperative Educational Laboratory

AC58 1219
Bilingual Early Childhood Education Learning System
Mrs. Shari Nedler
Southwest Educational Development Laboratory

AC59 1415
Multicultural Social Educational Program
Mrs. Martha Smith
Southwest Educational Development Laboratory

AC60 1410
Research and Development on Preschool Disadvantaged Children 5-1181
Professor Merle B. Karnes
Institute for Research on Exceptional Children

AC62 1148
Development of Materials for a One-Year Course in African Music for the General Undergraduate Student 6-1779
Dr. Vada E. Putcher, Principal Investigator
Howard University

2. *Products Mostly for Individualized Instruction

AC01 *2X37
Job Experience Kits 7-0111
Dr. John Krumboltz, Principal Investigator
Stanford University

AC09 2X45
Multi-Media Economics Curriculum Development Project 8-0447
Edmund W. Fitzpatrick, Principal Investigator
Educational Technology Center

AC10 2X33
Intermediate Science Curriculum Project (ISCS) 6-1762
Ernest Burkman, Principal Investigator
Florida State University

AC25 2X59
MATCH Box 5-0710
Mr. Frederick H. Kresse, Principal Investigator
The Children's Museum

AC41 2X20
Model CM Classrooms for Individualized Instruction - Grades 4-6
John C. Maxwell, Principal Investigator
Upper Midwest Educational Laboratory

AC47 2X37
Self-Instructional System in Basic Electricity
Lawrence D. Fish, Principal Investigator
Northwest Regional Educational Laboratory

AC50 2X37
Self-Instructional System in Welding
Lawrence D. Fish, Principal Investigator
Northwest Regional Educational Laboratory

3. *Products Mostly for Group Instruction

AC04 *3X33
Harvard Project Physics 5-1038
Professor Gerard Holton, Associate Professor
F. James Rutherford, Professor Fletcher G. Watson,
Principal Investigators
Harvard University

AC08 3535
Inquiry Materials for Social Studies HS-041 and H-292
Edwin Fenton, Principal Investigator
Carnegie-Mellon University

AC11 3X35
Preparation and Evaluation of Social Studies Curriculum Guides and Material for Grades K-14 5-0659
Edith West, Principal Investigator
University of Minnesota

AC12 3X55
The Taba Curriculum Development Project in Social Studies 5-1314
Dr. Norman Wallen, Principal Investigator
San Francisco State College

AC15 3X55
A Study of Geo-Historical Structure for a Social Studies Curriculum 6-1195
Ridgway Shinn, Jr., Principal Investigator
Rhode Island College

AC17 3X25
The Development of Instructional Materials and Teaching Strategies on Race and Culture in American Life 8-0197
Dr. John S. Gibson, Principal Investigator
Tufts University

AC20 3X31
The Oregon Curriculum: A Sequential Program in English 5-0366
Albert R. Kitzhaber, Principal Investigator
University of Oregon

AC23 3634
Unified Mathematics Program 7-0711
Dr. Howard F. Fehr
Columbia Teachers College

Products on Group Instruction (continued)

AC53 3X58
An Enlarged Music Repertory for Kindergarten Through Grade Six
5-0219
Gordon Hardy, Principal Investigator
The Juilliard School

AC56 3940
Western Interstate Commission for Higher Education, Planning, Analysis and Management Systems Project 8-0708
Dr. Robert Huff, Principal Investigator
WICHE/MPS

AC67 3X11R
Communication Skills Program
Robert W. O'Hare, Principal Investigator
Southwest Regional Laboratory for Educational Research and Development

AC52 3X37
A Junior High School Industrial Technology Curriculum Project
Industrial Arts Curriculum Project 7-0003
Donald G. Lux and Willis E. Ray Principal Investigators
The Ohio State University

4. *Products Mostly for Teacher Training

AC19 *4X40
Simulation Training in Planning Vocational Educational Programs and Facilities 7-0158 (M4)
Dr. Darrell L. Ward, Principal Investigator
Ohio State University

AC26 4740
Behavioral Objectives Package
Dr. James L. Olivero and Dr. Carmen R. Timiraos, Principal Investigators
Southwestern Cooperative Educational Laboratory

AC27 4X41
Backgrounds in Language
Mrs. Barbara K. Long, Principal Investigator
Upper Midwest Regional Educational Laboratory

AC28 4X44R
Individualizing Instruction in Mathematics (Minicourse 5)
Dr. Meredith D. Gall, Principal Investigator
Far West Laboratory for Educational Research and Development

AC33 4449
Cooperative Urban Teacher Education
Dr. Grant Clothier, Principal Investigator
Mid-Continent Regional Educational Laboratory

AC35 4X40
Instructional System in Development of Higher Level Thinking Abilities
Lawrence D. Fish, Principal Investigator
Northwest Regional Educational Laboratory

AC36 4X41
Dialects and Dialect Learning - An English Inservice Program
Karen Matison Hess, Principal Investigator
Upper Midwest Regional Educational Laboratory

AC38 4X40
Individualized Instruction Through Contingency Management
Mr. Lanny E. Morreau, Principal Investigator
Upper Midwest Regional Educational Laboratory

Products on Teacher Training (continued)

AC39 4X40
Instructional System in Facilitating Inquiry in the Classroom
Lawrence D. Fish, Principal Investigator
Northwest Regional Educational Laboratory

AC48 4X40
Instructional System in Systematic and Objective Analysis of Instruction
Lawrence D. Fish, Principal Investigator
Northwest Regional Educational Laboratory

AC49 4X40
Instructional System in Interaction Analysis
Lawrence D. Fish, Principal Investigator
Northwest Regional Educational Laboratory

AC63 4440
A Program for Leadership Training in Team Teaching
L. Jean York, Principal Investigator
University of Texas (Austin)

AC64 4X43
The Teaching of Science: A Self-Directed Personalized Teacher Education Program
David Butts and Gene Hall, Co-Investigators
University of Texas (Austin)

AC65 4940
Comprehensive Personal Assessment and Counseling Feedback Systems for Pre-Service Teacher Education Programs
Oliver H. Brown, Principal Investigator
University of Texas (Austin)

AC68 4940
Instruments and Procedures for Describing Effective Teacher Behavior
Leland L. Medsker, Principal Investigator
University of California (Berkeley)

5. *Products Mostly for Administering Programs

AC02 *5X57
The Vocational Development Inventory 5-0038
Dr. John O. Crites, Principal Investigator
University of Iowa

AC05 5X5X
A System for Individualizing and Optimizing Learning Through Computer Management of the Educational Process 8-0157
Dr. Alexander Schure, Principal Investigator
New York Institute of Technology

AC22 541X
Individual Readiness Test 9-F-017
Wanda Walker, Principal Investigator
Northwest Missouri State College

AC43 5519
Parent/Child Toy Lending Library
Glen Nimnicht, Edna Brown, Stan Johnson, Bertha Addison, Principal Investigators
Far West Laboratory for Educational Research and Development

AC61 5X33
Exploration in Biology Topics (Inquiry Skills Program)
Eugenia M. Koos
Mid-Continent Regional Educational Laboratory

AC66 5959
Alternatives for Learning Through Educational Research and Technology (ALERT): An Educational Information System
C. L. Hutchins, Principal Investigator
Far West Laboratory for Educational Research and Development

AC69 5940
An Instrument and Procedures for Improving Communication and Academic Policy Making
Leland L. Medsker, Principal Investigator
University of California (Berkeley)

6. *Products for Combined Programs

AC24 *6519
Educational Television for Preschoolers 8-0475 (Sesame Street)
Mrs. Joan Ganz, Principal Investigator
Children's Television Workshop

AC42 6940R
Effective Questioning - Elementary Level (Minicourse 1)
Dr. Walter R. Borg, Principal Investigator
Far West Laboratory for Educational Research and Development

AC70 6550R
Multi-Unit Elementary Schools
Dr. Herbert J. Klausmeier, Principal Investigator
Wisconsin R & D Center for Cognitive Learning

PROJECT DOSSIER

Title _____

Accession # _____ Classification # _____

Product logged on _____ by _____

Team assigned _____

_____

ERIC Search I result _____, on _____

NA

☐ Material requested from Principal Investigator on

_____ by _____

☐ Material received on _____ by _____

☐ Product retained in 1971 pool: ☐ Yes No ☐ by ____

☐ Precis drafted on _____ by _____

☐ Precis approved on _____ by _____

☐ Selected by Panel for extended study: ☐ Yes No ☐

☐ Feedback requested from Princ. Inves. on ____ by ____

☐ Feedback received on _____ by _____

From _____

Principal Investigator _____

☐ Extended analysis done on _____ by _____

☐ Site Visit made on _____ by _____

☐ Final analysis reviewed on _____ by _____

☐ Recommended by Panel for VP '71 _____

☐ Product description completed on ____ by ____

☐ Product description edited on ____ by ____

☐ Product description sent to NCEC on ____ by ____

☐ Dossier ready for ERIC-TM on _____ by _____

Notes:

As step is completed, date and initial.

Check NA only if statement is not applicable to product.

PROJECT LOG

Title _____

Accession # _____ Classification # _____

Product logged on _____ by _____

Team assigned _____

_____

ERIC Search I result _____, on _____

NA

☐ Material requested from Principal Investigator on

_____ by _____

☐ Material received on _____ by _____

☐ Product retained in 1971 pool: ☐ Yes No ☐ by ____

☐ Precis drafted on _____ by _____

☐ Precis approved on _____ by _____

☐ Selected by Panel for extended study: ☐ Yes No ☐

☐ Feedback requested from Princ. Inves. on ____ by _____

☐ Feedback received on _____ by _____

From _____

Principal Investigator _____

☐ Extended analysis done on _____ by _____

☐ Site Visit made on _____ by _____

☐ Final analysis reviewed on _____ by _____

☐ Recommended by Panel for VP '71 _____

☐ Product description completed on _____ by _____

☐ Product description edited on _____ by _____

☐ Product description sent to NCEC on ____ by ____

☐ Dossier ready for ERIC-TM on _____ by _____

Notes:

As step is completed, date and initial.

Check NA only if statement is not applicable to product.

Appendix E

Product Precis

	Date	Initial
1st Draft	_____	_____
Revision 1	_____	_____
Revision 2	_____	_____
Revision 3	_____	_____

Accession No. _____

Product Identification

OE Number _____

Title _____

_____

Principal Investigator: _____

Address _____

_____

_____

Product Classification

Code _____

Major Emphasis _____

Target Population _____

Age level _____

Brief Description of Product

Accession No. _____ Product Precis

Evaluation/Validation

Accession No. _____ Product Precis

Implementation Implications

Major requirement for implementation

Personnel requirements

Materials and Facilities

Other Administrative Considerations

Other Factors on limitations

Product Dissemination Efforts

Existing

Planned

Description of Obstacles to Implementation and Suggested Dissemination Strategies

Appendix E-5

Guidelines for Product Evaluation Precis

General Note:

The design of the 4 pages of the forms is intended to give you room to write and revise. The space allowed for the different headings was arbitrary and does not imply that you should fill it all. Be as brief as possible, without sacrificing completeness of important information. The panel must deal with 60 of these in a two-day meeting and must be able to make comparisons.

Page 1

Accession No.
Project Identification
Classification
Principal Investigator

These will be filled in before you get the dossier

Brief Description of Product

Most of this should come from the descriptive information supplied by the developer but may need some supplementation. It should include:

What it is

What it does-purpose and objectives, major characteristics, benefits claimed, and, if possible, how it differs from other products or programs available.

Who it is designed to serve - the group or groups for which it was developed and any limitations within this. (e.g. but not suitable for poor readers or only useful for high ability, college bound students)

Page 2 Verification - Evaluation/Validation

This area is particularly important and may eventually require the most additional information. In the first precis, this section should be primarily descriptive with analysis of the adequacy of the evaluation to be done if the product survives the first round of examination by the advisory panel.

Promotional brochures should be treated with caution. The kinds of information needed are:

Description of verification design and data, including such things as range of evaluation or validation - such as:

questionnaires to teachers, students, who, how many,
formal try-out design
breadth and size of sample - scope of try-out
objectives of the evaluative procedure

Indicate what reports on validation are available - either already in dossier or asked for from principal investigator

Page 3 Implementation Implications

Major requirements for implementation - Leave this blank initially - it will be filled in after panel consideration.

Fill in specific implementation requirements under headings - if none, simply write none.

Personnel required: number, special training, availability of special training

Material and facilities

Required, optional

Availability - ease, source

Cost - initial (i.e. one time)
ongoing (per pupil if possible)

Administrative Considerations

Organizational implications

Conditions for installation:

minimum feasible scope for try out-individual, class, department, whole school, school system

minimum feasible scope for adoption -

Other limitations or factors - anything else you turn up that would or should influence a decision to adopt, that isn't covered under other headings. (include anything you identify that developer hasn't mentioned but indicate that it is your addition by marking Reviewer Comment)

Guidelines

Page 4

Dissemination Efforts by Developer or Publisher

Existing - what has been done or is underway
Planned - what he expects to do in the future

The reason for this item, I assume, is to determine whether the product needs NCEC efforts or is already being widely disseminated. The answer to this may need to come from questions to the PI although in some cases of already widely used materials, it may be part of a report, brochure, or article.

Obstacles to implementaion and suggestsions for dissemination strategy

This is not so much a basic part of the precis as it is a required element of the final description.

This can be left blank initially unless you have something that occurs to you to include. The material for this will probably come out of advisory panel discussion.

Guidelines

Page 4

Dissemination Efforts by Developer or Publisher

Existing - what has been done or is underway
Planned - what he expects to do in the future

The reason for this item, I assume, is to determine whether the product needs NCEC efforts or is already being widely disseminated. The answer to this may need to come from questions to the PI although in some cases of already widely used materials, it may be part of a report, brochure, or article.

Obstacles to implementaion and suggestsions for dissemination strategy

This is not so much a basic part of the precis as it is a required element of the final description.

This can be left blank initially unless you have something that occurs to you to include. The material for this will probably come out of advisory panel discussion.

AC_____

Appendix F
P-J: 717-01

PEP

Detailed Product Analysis

☐ Site Visitor ☐ Product Reviewer

1. Goals/Effectiveness

2. Costs/Implementation

3. Staff Comments

Appendix G
DRAFT FOR COMMENT

Outline of Possible Criteria for Comparative Evaluation of Educational Products

At its initial meeting, the Appraisal Panel for the evaluation program will need to formulate the criteria by which products will be evaluated. This step is necessary to give further direction to the staff support work on product descriptions. This outline is intended to serve as a basis for discussions leading to the defining of evaluative criteria. It does not attempt to suggest procedures for utilizing the ratings in reaching a final decision about products, although these matters will also need to be determined by the panel at its initial meeting. In preparing the outline, an effort has been made to identify major characteristics of acceptable programs, to avoid as far as possible overlapping between characteristics chosen for evaluation, and to keep the number of characteristics relatively small.

Possible Criteria for Comparative Evaluation

1. Importance of Goals Sought

   ... How urgent are the tasks to which the product is addressed?

   ... Does content, method, or both differ markedly from conventional programs in the same field?

   ... To what extent is the product concerned with major goals and outcomes of education?

   ... How large is the group of students for whom the product may be considered appropriate?

Considering urgency, innovative features, concern with highly significant goals, and potential scope of application, which of the following best describes the importance of the product's goals?

4	Major
3	Substantial
2	Modest
1	Trivial

Comment: The evaluation of this characteristic calls for a subjective judgment based on complex underlying criteria. It may well be desirable to break this characteristic into separate aspects.

2. Evidence of Effectiveness

... How effective is the product in producing desirable educational outcomes?

... Are the empirical studies of effectiveness well designed? Adequate in scope? Based on comparative performance? Realistic?

Considering both the size of the effects on student behavior and the scope and quality of the empirical studies on which the evaluation of effectiveness is based, which of the following best describes the evidence on student performance?

4 Adequate evidence that the product produces substantial effects on student behavior

3 Adequate evidence that the product produces modest effects on student behavior

2 Preliminary or otherwise inconclusive evidence available on effects

1 Little or no empirical evidence on effectiveness available

Comment: The panel may need to decide whether evidence of effectiveness should be conceived narrowly in terms of the product's specific objectives or more broadly in terms of general outcomes.

3. Evidence of Educational Soundness

... Was product designed on the basis of clearly formulated objectives stated in behavioral terms?

... Were objectives subjected to external challenge during the development process?

... Is content appropriate to the objectives?

... How well organized is the product with respect to scope and sequence of activities?

... To what extent is methodology designed in terms of the goals of instruction?

... Is appropriate flexibility in methodology provided?

... Have aids to evaluation been developed specifically to be appropriate to the objectives, content, and methods of the program?

Considering the clarity and adequacy of the objectives, the appropriateness of the content and methods, and the provisions for suitable evaluation, which of the following best describes the intrinsic evidence of the product's educational soundness?

4 Clear evidence of thorough developmental efforts on all major aspects

3 Product basically sound but somewhat deficient in one major aspect of development

2 Product development barely adequate, on the whole

1 Product shows serious deficiencies in development which make its soundness debatable or evidence regarding development is inadequate to permit a judgment on the product.

Comment: This characteristic might be more useful if broken into separate aspects.

4. Readiness for Adoption

... Has product been field-tested under realistic conditions for feasibility and acceptability?

... Have all needed components of the product been completed for use by adopting schools?

... Have necessary planning data been developed explicitly to aid schools on reaching a decision concerning adoption?

... Is material readily available?

Considering evidence on field-testing, on completeness of package and availability of planning data, how would you describe the product?

4 Ready to be considered widely for adoption

3 All basic components complete; could be packaged for adoption with a small amount of additional work

2 Could be offered for adoption despite significant lacks in package from user's viewpoint

1 Has not been subjected to appropriate field testing or requires substantial further development of components.

5. Costs

... Would large per-student costs be incurred in acquiring equipment and supplies?

... Would additional professional staff be required?

... Would substantial in-service training be required for teachers?

... Would the services of outside consultants be required?

... Would substantial additional space be required?

... Would introduction require substantial modification of school organization, schedule and/or administrative procedures?

... Does product require substantial additional time for teachers, students or both on a continuing basis <u>after</u> the initial breaking-in period?

... Does product require substantial additional supplies and supporting facilities on a continuing basis after the initial breaking-in period?

... Is the product likely to arouse significant negative attitudes on the part of students, teachers, parents or other members of the community?

Giving chief weight to the probable dollar costs, but taking account of other significant costs, which of the following best describes the effect on school costs of adopting the product?:

1 Would represent a major commitment of school resources

2 Would represent a substantial commitment of school resources

3 Would represent a modest commitment of school resources

4 Would be no more costly (or possibly less costly) than traditional ways of attaining similar objectives.

<u>Comment</u>: The handling of the cost aspect in evaluating products for national dissemination is likely to be particularly difficult. In particular, it may be desirable to distinguish initial from continuing costs.

ERIC-TM

Product Evaluation Project

Formulation of Criteria

Introduction

At its initial meeting on March 3 and 4, 1971, the Panel devoted most of its effort to the formulation of a set of criteria with which to describe and evaluate products, and a plan for applying these criteria to the task of evaluating the 1971 Product Pool. Most of the discussion which follows is organized around four clusters of criterion-related elements: goals, effectiveness, "costs," and adoptability. Toward the end of the paper, the basis for an overall judgment is shown. This encompasses not only the four clusters, but also USOE priorities and other elements into a single "global" determination. A combination of check lists and rating scales has been suggested as an integral part of criteria application. These are described at the appropriate points in the discussion.

The Rating Form - Product Evaluation Pool '71, attached, is a work sheet the Panel proposes to use for: (a) the initial individual rating of each product, (b) recording the results of group consideration of each product during the initial reading period, (c) making a second sequence of judgments on each product retained in the pool, after the detailed analyses on this group have been completed, (d) confirming the judgments related to products set aside, and (e) monitoring the internal characteristics of the rating process. It is suggested that, in reviewing this document, the reader step through the Rating Form as he proceeds through the clusters in the paper. This procedure will be found helpful in clarifying the relationship between the criteria formulation and the plan for its application to the evaluation task.

In considering the criteria, the following working assumptions of the Panel should be kept in mind:

(1) The product evaluations are intended to describe the relative importance of dissemination of the product

as of the spring of 1971. It is recognized that decisions concerning products developed in the current evaluation will need to be reviewed in subsequent years not only because of changes in the product but also because of changed conditions and priorities in American education.

(2) The descriptions and evaluations of products with respect to specific criteria are designed to aid each panel member in arriving at an overall judgmental evaluation of a given product rather than eliciting from him ratings on components to be used in numerical calculations.

(3) Evaluations given to a product by different panel members will be combined by discussions leading to a consensus -- where consensus is possible -- rather than by a numerical averaging process.

(4) It will be possible, after the fact, to reconstitute an overall rating on different grounds by holding the ratings on clusters by panel members stable and reflecting changed outlooks regarding USOE priorities, dissemination and other similar considerations.

Criteria

Goals: Extent to which the product may be expected to have major effects on significant educational outcomes. Four separate aspects of this criterion are to be considered:

(1) Urgent Present Need - Does the product address itself to urgent needs?

(2) Desirable Originality - Does the product embody well-conceived innovations in content, method or both? Does it reflect old orientations or new ones?

(3) Educational Centrality - Do the goals represent a broad spectrum of outcomes or narrow ones? Does the product concern itself with major outcomes or minor ones? Is product content appropriate to the stated goals?

(4) Size of Target Population - How large is the group for whom the developer considers the product appropriate?

Raters will designate with contrasting symbols any of the four aspects on which a product is notably strong or notably deficient. If some other aspect of the product's goals deserve comment, the rater is asked to describe this aspect briefly and to designate whether it is a strong or weak point of the product. The summary rating for this criterion will be expressed on the following scale:

A Excellent

B Good

C Possible

D Unimpressive

A plus sign may be added to the rating if the rater wishes to do so.

Provision is made for the possibility that the rater will wish to evaluate the product under different assumptions concerning the products goals. If so, he is asked to show his alternative evaluation on a separate rating sheet. This option would be useful, for example, if the author's claimed goals for the product were inconsistent with the product as developed, in the judgment of the rater.

Effectiveness: Extent to which the product is effective in accomplishing its stated goals and in accomplishing other goals than those stated by its developer (side effects). Judgments on this criterion are to be based on evidence available in Spring, 1971. Three aspects of effectiveness are to be considered:

(1) Adequacy of Test Data - How adequate are test data with respect to sample size, fairness of sample with respect to target population, provision of data for appropriate control groups, and objectivity of judgments about the product.

(2) Performance in Field Trials - How well did the

product perform in field trials? Is there evidence that teachers and students accepted the product readily and wish to continue to use it?

(3) Internal and Background Evidence - Is there internal evidence of product quality or evidence on its background (e.g., previous outstanding performance of the developers in producing highly effective products) which offers useful supplementary indications of the product's probable success or failure?

With respect to side effects, the rater will state briefly the particular unanticipated outcomes with which his rating is concerned. For example, a programmed text in algebra might have produced gains in reading ability. On the other hand, adverse emotional effects might arise from certain ways of teaching reading. The observations of the raters with respect to side effects will be used to seek further evidence concerning them when site visits are made.

Separate ratings on effectiveness will be made with respect both to stated goals and side effects. In both instances, the following scale will be used:

A Excellent

B Good

C Fair

D Unimpressive

F Undesirable

? Impossible to estimate on the basis of existing evidence

Here again, a plus sign may be added to either rating if the rater wishes to do so.

Costs: Extent to which the introduction and subsequent use of the product place heavy demands on the school's resources, both economic and human. In evaluating the costs of a product, the rater is asked to consider both financial outlay and a cluster

of nine other aspects as follows:

(1) Increase or decrease in professional staff.

(2) Increase or decrease in maintenance staff.

(3) Inservice training costs in dollars and time.

(4) Consultant costs related to adoption.

(5) Repair and substitute costs.

(6) Costs for extra space.

(7) Increase or decrease in time needed by students, teachers, administrators.

(8) Disruption of the system.

(9) Opposition by community, students, and staff.

The last three factors will receive special consideration since products which can easily be introduced without disturbing the on-going system and products which can be introduced without extensive involvement of higher-level administrators are particularly likely to repay dissemination efforts. On the other hand, dissemination of products which, for example, seem to teachers to downgrade their professional role, might be expected to encounter serious resistance.

Separate ratings on costs will be made with respect both to installation and maintenance. For installation, the three levels are: Negligible, Modest, and High. Space is provided for recording a dollar figure for installation's financial outlay. For maintenance, the three levels are related to annual per pupil costs for the more traditional ways of achieving the same objectives. The three levels, with rough dollar guidelines for each level, are as follows: Normal Range ($0 to $4), High ($4 to $14) and Very High (More than $14). For total installation and maintenance costs, the rater may indicate that the existing evidence does not permit him to make a judgment(?) Space is also provided to record the rater's judgment of an estimated initial installation cost and of an estimated per-pupil annual cost.

Adoptability: Extent to which the product is readily available to a school which wishes to adopt it. Three levels related to materials have been specified, as follows:

(1) Immediate, virtually unlimited availability

(2) Immediate availability, but supply is limited

(3) Only sample materials presently available

A fourth aspect of the cluster covers availability of materials for use in product testing and evaluation.

Three significant administrative constraints affecting adoptability also have been specified, as follows:

(1) Requires special training in advance of use

(2) Likely to be system-disruptive

(3) Plant installation required

Space is provided to indicate lead-time, in years, for installation.

A rating scale has not been adopted for use for the ADOPTABILITY criterion. Rather, the materials and administration related factors of relevance are identified. Thus, raters in this case apply their judgment by checking any of the factors that apply.

Overall Judgment of Need for USOE 1971 Dissemination Support: Extent to which the product deserves favorable consideration by the panel for inclusion in the recommended group of products. Judgment is to be based on the most up-to-the-minute data available, and is to encompass in one global estimate the separate judgments made regarding goals, effectiveness, "costs," and adoptability.

In addition to epitomizing the specific criterion-based evaluations, the overall rating will take account of the following:

(1) Concurrence of product with USOE priorities as listed in Attachment A.

(2) Possible effects of NCEC dissemination efforts considered in the light of existing support (if any) for dissemination. Are sufficient dissemination efforts already being made? Would this dissemination support provide needed impetus to get the product over the hump?

(3) Impact or leverage or multiplication effect of support. This includes size of market, visibility of results, and breadth of goals sought. Would dissemination support accelerate implementation to a significant extent? Would it accelerate school improvement?

(4) Existence of alternative products, comparable in adoptability and at least equally cost-effective and desirable. Does this product introduce something unique? Are there equally good products to serve the same need as effectively and in the same range of "costs"? Space is provided for designating critically competitive products.

The rater will make notations on the aspects of which he has taken account, identify critically competitive products and then subsume all the component judgments into a final rating.

Ratings for the global criterion will be expressed on a five step scale as follows:

A Excellent

B Good

C Fair

D Unimpressive

F Undesirable

A plus sign may be added to the rating if the rater wishes to do so, in effect making a nine-step scale.

Educational Testing Service
Princeton, New Jersey

March 12, 1971

Attachment A

Considerations in the Establishment of ETS Evaluation Criteria

. Office of Education priorities

1.1 Present priorities

1.1.1 Right to Read

1.1.2 Disadvantaged - elementary and secondary

1.1.3 Equal Educational Opportunity

1.1.4 Environmental/Ecological Education

1.1.5 National Institute of Education

1.1.6 Experimental Schools

1.1.7 Disadvantaged - post secondary

1.2 Possible future priorities

1.2.1 Create opportunities through education of the handicapped

1.2.2 Stimulate development of innovative and effective approaches to education

1.2.3 Meet the needs of economically disadvantaged children

1.2.4 Eliminate racial, ethnic and cultural barriers to educational opportunities

1.2.5 Stimulate career education programs

RATING FORM — PRODUCT EVALUATION POOL '71 — ETS

---DRAFT VI---

PRODUCT CODE NUMBER: AC _____
RATER'S NAME / CODE: .....

(Further descriptions if desired; Aim, or Author, Institutional Sponsor etc. .....

Date of Rating / /71

GOALS

[A = Excellent B = Good C = Possible D = Unimpressive]

Considerations: Urgent Present Need / Desirable Originality / Educational Centrality / Size of Target Population / ..... (other)

Circle or underscore those of unusual positive significance.
Strike out any that struck you as notably deficient.

Rating: A B C D (Circle one or more; or add ± if preferred)

Indicate here if alternative conception of goals or target etc. has been used as a basis for second rating sheet ☐

EFFECTIVENESS

Using data of 1971

[A, B, C, D, or "F" = Undesirable, or "?" = No reasonable estimate possible at this time]

Considerations: Adequacy of Test Data (Sample Size; Sample Fairness; Appropriate Controls; Objectivity of Judgment) / Performance in Field Trials Including Teacher Feed-back / Background or Internal Evidence Suggesting Success or Failure e.g. Prior Track Record

Rating on Stated Goals: A B C D F ?

Rating on Side Effects or Unanticipated Effects: A B C D F ? ..... (Describe Side Effect)

"COSTS"

To User

[If you average costs across students, indicate how many .....
If you average costs across years, indicate how many .....]

Considerations: Extra (or less) Professional Staff / Extra (or less) Maintenance Staff / In-service Training Costs / Consultant Costs / Repair and Substitute Costs / Extra Space Costs / Extra (or less) Time Required / System Disruption / Opposition ·Students ·Staff ·Community

Rating:

	Negligible	Modest	High	?
Installation:	$0-4	$4-14	>$14	Student/year
"Maintenance":	Normal Range	High	Very High	?

+$.....(one-shot)
OR SAVINGS (estimated)
+$..... (per annum)

ADOPTABILITY

Materials: Immediate Unlimited Availability ☐ / Immediate Limited Availability ☐ / Sample Materials Only ☐ / Test and Evaluation Materials Availability ☐

Check Relevant Boxes

Administration: Requires special training in advance of use ☐
Likely to be system-disruptive ☐
Plant installation required ☐ (lead-time.....years)

OVERALL

Considerations Circle factors, if any, that particularly affected your judgment

Above Judgments About: Goals Effectiveness / "Costs" Adoptability / USOE Priorities / Other Support Available (or not available) / "Impact" or "Leverage" or "Multiplication" Effect of Support / Existence of Alternative Products comparable in adoptability and at least equally cost-effective and desirable

[CRITICALLY COMPETITIVE PRODUCT(S) _____]

Rating of need for USOE 1971 dissemination support based on data available 2/71 — 3/71

A B C D F

ERIC-TM

Product Evaluation Project

Formulation of Criteria

Introduction

At its initial meeting on March 3 and 4, 1971, the Panel devoted most of its effort to the formulation of a set of criteria with which to describe and evaluate products, and a plan for applying these criteria to the task of evaluating the 1971 Product Pool. Most of the discussion which follows is organized around four clusters of criterion-related elements: goals, effectiveness, "costs," and adoptability. Toward the end of the paper, the basis for an overall judgment is shown. This encompasses not only the four clusters, but also USOE priorities and other elements into a single "global" determination. A combination of check lists and rating scales has been suggested as an integral part of criteria application. These are described at the appropriate points in the discussion.

The Rating Form - Product Evaluation Pool '71, attached, is a work sheet the Panel proposes to use for: (a) the initial individual rating of each product, (b) recording the results of group consideration of each product during the initial reading period, (c) making a second sequence of judgments on each product retained in the pool, after the detailed analyses on this group have been completed, (d) confirming the judgments related to products set aside, and (e) monitoring the internal characteristics of the rating process. It is suggested that, in reviewing this document, the reader step through the Rating Form as he proceeds through the clusters in the paper. This procedure will be found helpful in clarifying the relationship between the criteria formulation and the plan for its application to the evaluation task.

In considering the criteria, the following working assumptions of the Panel should be kept in mind:

(1) The product evaluations are intended to describe the relative importance of dissemination of the product

as of the spring of 1971. It is recognized that decisions concerning products developed in the current evaluation will need to be reviewed in subsequent years not only because of changes in the product but also because of changed conditions and priorities in American education.

(2) The descriptions and evaluations of products with respect to specific criteria are designed to aid each panel member in arriving at an overall judgmental evaluation of a given product rather than eliciting from him ratings on components to be used in numerical calculations.

(3) Evaluations given to a product by different panel members will be combined by discussions leading to a consensus -- where consensus is possible -- rather than by a numerical averaging process.

Criteria

Goals: Extent to which the product may be expected to have major effects on significant educational outcomes. Four separate aspects of this criterion are to be considered:

(1) Urgent Present Need - Does the product address itself to urgent needs?

(2) Desirable Originality - Does the product embody well-conceived innovations in content, method or both? Does it reflect old orientations or new ones?

(3) Educational Centrality - Does the product concern itself with outcomes that are central to education rather than with special outcomes?

(4) Size of Target Population - How large is the group for whom the developer considers the product appropriate?

Raters will designate with contrasting symbols any of the four aspects on which a product is notably strong or notably deficient. If some other aspect of the product's goals deserve comment, the rater is asked to describe this aspect briefly and to designate whether it is a strong or weak point of the product. The summary rating for this criterion will be expressed on the following scale:

A Excellent

B Good

C Possible

D Unimpressive

A plus or minus sign may be added to the rating if the rater wishes to do so.

Provision is made for the possibility that the rater will wish to evaluate the product under different assumptions concerning the products goals. If so, he is asked to show his alternative evaluation on a separate rating sheet. This option would be useful, for example, if the author's claimed goals for the product were inconsistent with the product as developed, in the judgment of the rater.

Effectiveness: Extent to which the product is effective in accomplishing its stated goals in its target population and in accomplishing goals other than those stated by its developer or producing outcomes in populations other than its target population (side effects). Judgments on this criterion are to be based on evidence available in Spring, 1971. Three aspects of effectiveness are to be considered:

(1) Adequacy of Test Data - How adequate are test data with respect to sample size, fairness of sample with respect to target population, provision of data for appropriate control groups, and objectivity of judgments about the product.

(2) Performance in Field Trials - How well did the

product perform in field trials? Is there evidence that teachers and students accepted the product readily and wish to continue to use it?

(3) Internal and Background Evidence - Is there internal evidence of product quality or evidence on its background (e.g., previous outstanding performance of the developers in producing highly effective products) which offers useful supplementary indications of the product's probable success or failure? Is product content appropriate to stated goals?

With respect to side effects, the rater will state briefly the alternative goals, alternative populations, or unanticipated outcomes with which his rating is concerned. For example, a programmed text in algebra might have produced gains in reading ability. On the other hand adverse emotional effects might arise from certain ways of teaching reading. The observations of the raters with respect to side effects will be used to seek further evidence concerning them when site visits are made.

Separate ratings on effectiveness will be made with respect both to stated goals and side effects. For stated goals, the following scale will be used:

A Excellent

B Good

C Fair

D Unimpressive

F Undesirable

? Impossible to estimate on the basis of existing evidence

For side effects, the interpretation of certain letter ratings may be modified as follows:

C Neutral

D Somewhat negative

F Very negative

Here again, a plus or minus sign may be added to either rating if the rater wishes to do so.

Costs: Extent to which the introduction and subsequent use of the product place heavy demands on the school's resources, both economic and human. In evaluating the costs of a product, the rater is asked to consider both financial outlay and a cluster of ten other aspects as follows:

(1) Materials

(2) Increase or decrease in professional and/or paraprofessional staff

(3) Increase or decrease in technical staff (audio-visual, computer, etc.)

(4) Inservice training costs in dollars and time

(5) Consultant costs related to adoption and maintenance

(6) Repair and substitute costs

(7) Costs for extra space

(8) Increase or decrease in time needed by students, teachers, administrators

(9) Disruption of the system

(10) Opposition by community, students, and staff.

The last three factors will receive special consideration since products which can easily be introduced without disturbing the on-going system and products which can be introduced without extensive involvement of higher-level administrators are particularly likely to repay dissemination efforts. On the other hand, dissemination of products which, for example, seem to teachers to downgrade their professional role, might be expected to encounter serious resistance.

Separate ratings on costs will be made with respect to installation, maintenance, and non-dollar costs. For installation, the three levels are: Negligible, Modest, and High. For maintenance, the three levels, Normal Range, High, and Very High, are related to annual per pupil costs for the more traditional ways of achieving the same objectives. The following estimates for the annual maintenance cost of a single course may be used as guidelines: Normal Range ($0 to $4), High ($4 to $14), and Very High (More than $14). For non-dollar costs, the three levels are Minimal, Tolerable, Excessive. For total installation, maintenance and non-dollar costs, the rater may indicate that the existing evidence does not permit him to make a judgment by using "?" as his rating. Space is provided to record the rater's judgment of estimated costs or savings.

Adoptability: Extent to which the product is readily available to a school which wishes to adopt it. Four considerations related to materials have been specified as follows:

(1) Immediate, virtually unlimited availability

(2) Immediate availability, but supply is limited

(3) Only sample materials currently available

(4) Materials for use in product testing and evaluation available

Three significant administrative considerations affecting adoptability also have been specified, as follows:

(1) Requires special training in advance of use

(2) Likely to be system-disruptive

(3) Plant installation required

Space is provided to indicate lead-time, in years, for installation.

Adoptability will be rated on the following scale:

A Easily adoptable

B

C Few problems

D

F Adoption too difficult

? Impossible to estimate on the basis of existing evidence

Overall Judgment of Need for USOE 1971 Dissemination Support: Extent to which the product deserves favorable consideration by the panel for inclusion in the recommended group of products. Judgment is to be based on the most up-to-the-minute data available, and is to encompass in one global estimate the separate judgments made regarding goals, effectiveness, "costs," and adoptability.

In addition to epitomizing the specific criterion-based evaluations, the overall rating will take account of the following:

(1) Concurrence of product with USOE priorities as listed in Attachment A.

(2) Possible effects of NCEC dissemination efforts considered in the light of existing support (if any) for dissemination. Are sufficient dissemination efforts already being made? Would this dissemination support provide needed impetus to get the product over the hump?

(3) Impact or leverage or multiplication effect of support. This includes size of market, visibility of results, and breadth of goals sought. Would dissemination support accelerate implementation to a significant extent? Would it accelerate school improvement?

(4) Existence of alternative products, comparable in adoptability and at least equally cost-effective and desirable. Does this product introduce something unique? Are there equally good products to serve the same need as effectively and in the same range of "costs"? Space is provided for designating critically competitive products.

Attachment A

Considerations in the Establishment of ETS Evaluation Criteria

1. Office of Education priorities

   1.1 Previous priorities

      1.1.1 Right to Read

      1.1.2 Disadvantaged - elementary and secondary

      1.1.3 Equal Educational Opportunity

      1.1.4 Environmental/Ecological Education

      1.1.5 National Institute of Education

      1.1.6 Experimental Schools

      1.1.7 Disadvantaged - post secondary

   1.2 Present priorities

      1.2.1 Create opportunities through education of the handicapped

      1.2.2 Stimulate development of innovative and effective approaches to education

      1.2.3 Meet the needs of economically disadvantaged children

      1.2.4 Eliminate racial, ethnic and cultural barriers to educational opportunities

      1.2.5 Stimulate career education programs

The rater will make notations on the aspects of which he has taken account, identify critically competitive products and then subsume all the component judgments into a final rating.

Ratings for the global criterion will be expressed on a five step scale as follows:

A Excellent

B Good

C Fair

D Unimpressive

F Undesirable

A plus sign may be added to the rating if the rater wishes to do so, in effect making a nine-step scale.

Educational Testing Service
Princeton, New Jersey

April 27, 1971

Appendix J

Suggested Procedures for Screening and Selection of Products for NCEC

General

The plan calls for identification of the more promising product prospects on the basis of an initial evaluation of all fifty products, closer study of each of those screened, and subsequent selection of a limited number of products to be recommended to NCEC for dissemination attention. Three alternative procedures generally applicable both for the initial March (24-25) and final (April 29-30) evaluations are suggested. For convenience, they are described in terms of the initial screening. Five questions permeate all three options. For your prior consideration, they are noted below.

1. Should there be one or more discussion sessions during the reading period?
2. Should subpanel members rotate during the day (e.g., after each set of "n" products)?
3. Should there be an opportunity for a subpanel to reopen an earlier decision after all rating is completed?
4. Should the subpanel have the option of recommending that the product be evaluated by another panel before a decision is reached?
5. Should an attempt be made to adjust ratings statistically to take account of leniency and severity of ratings?

Structuring the Rating and Evaluation Process

Each of the plans is based on the following assumptions:

1. Each panel member will evaluate twenty-five products,
2. Each product will be evaluated by four panel members,
3. The evaluation of each product will involve a detailed examination of the product in terms of the considerations included in the Rating Form,
4. The four panelists who evaluate a particular product will reach a consensus on the overall rating which they wish to assign to it.

Within this framework, and as suggested in the list of questions above, decisions are needed about the following issues in order to develop a working plan:

1. How reading periods and discussion periods are to be scheduled,
2. How subpanels are to be constituted,
3. How to take account of possible differences between subpanels in reaching a decision about a product.

For concreteness of discussion, three possible procedures are described.

Plan I - The Production Model. Each panelist works through his twenty-five products and rates each of them. He then joins three other panel members, all of whom have rated the same twenty-five products. The four-man team arrives at a consensus for each of their products. A small-scale statistical experi-

ment is conducted to develop an "adjustment factor" for each team to take account of possible differences between the two teams in rating standards.

Plan II - The Alternating Activities Model
This approach would differ from the production model in two ways. First, each panel would hold discussion sessions to arrive at consensus for the first four products, the next four products, the next eight products and the final nine products. This plan is based on the view that the 'discussion' sessions would be useful not only in arriving at consensus but would also clarify the evaluation process. Second, although no statistical calibration of the panels would be attempted, a panel could, if it were unable to reach a firm consensus on a product, recommend that the other panel be asked to evaluate that product before any final action on the product was taken. In addition, if it turned out that the general level of ratings was markedly higher for one subpanel than for the other, each panel would be asked to discuss possible causes of the difference, then to reconsider its ratings in the light of the observed differences. There would, of course, be no obligation for a panel to change its ratings.

Plan III - The Panel Rotation Model. This approach would differ from the Alternating Activities Model in one major respect. The assignment of products to raters would be so arranged that each panel member would be assigned to three different four-person teams. Every panelist would be on at least one team with every other panelist. Thus, during the course of the reading, he would have the opportunity to exchange views with every other panel member. (For mechanical reasons, each panelist would read twenty-five papers in common with one other member, nine papers in common with two other members and eight papers in common with each of the other four members.) This model would be somewhat more difficult to manage than the other two models. However, it should substantially reduce the danger of lack of comparability of ratings across panels, especially if each panel held a brief review meeting after seeing how its distribution of ratings compared with the distribution for the other five panels.

Allocation of Work Load Suggested random organization of subpanels for Plans I and II together with random allocation of products is shown on page 3. A comparable listing of subpanels and products for Plan III is shown on page 4. In both cases, the ordered list of products was used to assemble odd-ordered and even-ordered items: by taking every other item, first starting with the first item (AC16), then with the second (AC62). AC numbers are grouped to accord with the tabs in your books.

Group Convergence Under any of the plans, individual panelists would rate a block of products and complete a rating form for each of the products in that block. Rating sheets would be picked up by a recorder who would tally ratings on a chart for use during the discussion period. The chart would present overall ratings in block units; when all four panelists had rated all products in that block and the ratings were recorded, the group discussion would, theoretically, move the group toward convergence or even perhaps consensus regarding the overall rating of each product in the block. As individual products come up for discussion, the recorder would return rating forms to voters and replace them in the working file when the discussion was finished. A sketch of an imaginary tally chart is shown on Page 5.

PEP Reading & Rating Schedule for March 24-25, 1971

Under Plan I or II

Odd-Ordered Accessions

Panel A

R. Heinich*
R. Mancuso
M. Scriven
B. Watson

Tab 1	Tab 2
16	10
44	50
59	09
60	

Tab 3	Tab 4
08	63
56	65
17	36
04	64
12	19
53	48
52	38

Tab 5	Tab 6
22	70
69	
61	
05	

* K. Komoski Vice R. Heinich

Even-Ordered Accessions

Panel B

R. Gousha
D. Krathwohl
M. Molyneaux
R. Stake

Tab 1	Tab 2
62	41
58	47
45	01
	25

Tab 3	Tab 4
23	33
67	26
20	68
11	27
15	28
	39
	49
	35

Tab 5	Tab 6
43	24
66	42
02	

PEP Reading & Rating Schedule for March 24-25, 1971

Under Plan III

Odd-Ordered Accessions	Even-Ordered Accessions
SESSION 1 **Panel I** R. Heinich* R. Mancuso M. Scriven B. Watson Tab 1: 16, 44, 59, 60 Tab 2: 10, 50, 09 Tab 3: 08	**Panel II** R. Gousha D. Krathwohl M. Molyneaux R. Stake Tab 1: 62, 58, 45 Tab 2: 41, 47, 01, 25 Tab 3: 23
SESSION 2 **Panel III** R. Heinich R. Mancuso M. Molyneaux R. Stake Tab 3: 56, 17, 04, 12, 53, 52 Tab 4: 63, 65, 36	**Panel IV** R. Gousha D. Krathwohl M. Scriven B. Watson Tab 3: 67, 20, 11, 15 Tab 4: 33, 26, 68, 27
SESSION 3 **Panel V** R. Gousha R. Heinich D. Krathwohl R. Mancuso Tab 4: 64, 19, 48, 38 Tab 5: 22, 69, 61, 05 Tab 6: 70	**Panel VI** M. Molyneaux M. Scriven R. Stake B. Watson Tab 4: 28, 39, 49, 35 Tab 5: 43, 66, 02 Tab 6: 24, 42

* K. Komoski Vice R. Heinich

PEP TALLY CHART

Product Ratings

A	A- B+	B	B- C+	C	C- D+	D	D- F+	F
9	8	7	6	5	4	3	2	1

Rater / AC#	G	K	M	St
62	7	3	1	9
58	6	6	5	4
45	8	7	8	9
41	5	4	4	2
47	3	1	2	2
etc.		etc.	etc.	

If consensus were not possible and even convergence difficult, the product might be referred to jury action by the panel containing the other four panelists.

Post-Screening Activity. By some mutually agreeable means, the panel will identify a limited number of products (not more than thirty-four, hopefully not less than twenty). After identification of the screened products which remain for detailed analysis, a task force of evaluation specialists from EPIE will meet to complete independent reviews of the dossiers and precis for those products, to add their judgments to judgments made by rating panels concerning (a) emphases for detailed analyses and (b) special information to be sought during site visits. Insofar as possible, the task force will assume the role of devil's advocate, to introduce into the record the kinds of comment concerning precis and products especially regarding effectiveness, that might have come from product-competitors, had it been feasible to search that sector. This group also will make specific suggestions on strengthening precis prior to panel's final selection meeting.

By the time of the third meeting, all sites will have been visited, needed additional information turned up, if available, dossiers expanded and precis brought up to date to reflect the late April condition.

Selection Process. A modification of the screening process, using similar forms, schedules and procedures could be applied to the process of selection.

Educational Testing Service
Princeton, New Jersey
March 19, 1971

Appendix K

On Interpreting Panel Summaries

by Michael Scriven

Summaries of the discussions of a panel whose aim is rank-ordering a set of proposals should not be read as if they were individual appraisals without a context. For example, the summary might just say "weak on innovation." What this means is relatively weak on the innovation dimension (and by implication, adequate on all the others, whatever they might be). The unhappy proposer is likely to think of the fourteen lovingly detailed innovative practices in his proposal, a powerful combination indeed--enough to make his school system more innovative than any other in the U.S., "Weak on innovation," indeed! But the fact is that some of the other proposals--perhaps only one--were stronger on this and comparable on the other dimensions of comparison. Proposals are not about what is, but about what ought to be. It is indeed conceivable that no other proposals were stronger on innovation but the panel felt that--in the light of the RFP and contemporary knowledge--they should have been. A ranking does not imply satisfactory performance by the leaders.

The proposer who receives or sees such a summary will also no doubt think sadly or angrily of the 50 pages he devoted to an organization plan and timelines. "What sort of a panel is this," he asks, "to make no mention at all of such a careful plan?" But no news is good news. Ranking only requires discriminating between complexes; it does not require absolutistic evaluation of each component. If management is not mentioned, management was competitive--and that is likely to mean that it met, in its own individual way, the ideal standards of the panel. Considerable praise!

Panels frequently do more than rank-order. Often they also identify a cutting-point in the ranking, separating the proposals that they recommend for funding from the rest. For someone who does not receive funding--let us suppose that his proposal is the best of those not funded--it may seem that a more detailed justification is called for in this case than in the case of a ranking. Surely "Not innovative enough" is inadequate? More is involved in the cognitive processes of the committee, but it is not necessarily amenable to informative verbal formulation. It would simply look like this. "The judgment of global merit of this proposal, given the details of available funding and the likely costs of the higher-ranked proposals when modified as suggested, excluded the alternative of recommending sharply reduced funding of the higher proposals combined with some funding of this one." That kind of judgment may have required hours of discussion to achieve, discussion which in turn presupposed many complex perceptions of the proposal which all panelists shared (since they were never challenged). It cannot be adequately rendered in a brief verdict. In fact, because of its presuppositions, it is not fully-supported even by the full tape-recording. Incidentally, the words underlined in the last paragraph show why it is misleading to make inferences from the written form of a winner's proposal. It is also misleading to base very much on a site visit to the installed project, since many variations in it may become necessary under the pressure of changing political, economic, experiential and evaluative pressures.

Finally, one might ask whether there is not some, better way to convey the panel's judgment to the proposer--who has, after all put much of his work and creativity into it. For the reasons given above, verbal analysis or even taped recordings are misleading. The panel could be commissioned to do a detailed verbal appraisal of each proposal, but its perception of what common background it shares with the proposer and what needs to be filled in is not reliable. And the task would certainly double--perhaps triple--the time required by the panel to reach a ranking (and cutting score).

A much better procedure would involve a discussion by the panel and the proposer. For then the proposer can indicate an area of puzzlement, and the panel - whose complete transcript to that point may involve not one word on this issue - can explain their perception of it. One might even discover oversights that would lead to a re-ranking, by this procedure. It would approximately double the panel's time-commitment in Washington, however, and that usually means the panel could not be convened since conflicts increase exponentially with increases in panel time. (Individual reading of the proposals, prior to convening, can often be fitted into a schedule.)

A much weaker alternative would be discussion with a representative of the panel, or with a staff member present throughout the entire panel discussion. But this procedure would still be greatly preferable to an unresponsive brief summary. Summaries themselves are in fact undesirable because they have the same verbal term as isolated appraisals and hence arouse inappropriate expectations. A list of "pros & cons" is perhaps the best verbal form, with some indication of relative weighting where there are large inter-item differences. It must of course be read as listing discriminators not absolutist merits and one of its virtues is that it can concisely indicate tensions that were not resolved by the panel. Their presence does not indicate a contradictory stance, but the fact that the committee was of different minds on some points and did not find it necessary to resolve the issue in order to achieve overall closure.

In conclusion, at the least there is a difference in skills between good evaluating and good writing-up for the originator, and, in addition to the difference in time required for the double job, this means that the best evaluation panel will not be available for the double job.

Comments on Procedures for Panel Ratings of Products

by Robert E. Stake

In some circumstances an agency will designate a panel of persons to pass judgment on each of a number of products so as to rank them or classify them as to merit or to identify a pool of meritorious products. The procedures the panelists should follow will differ depending on the degree to which they are in communication with each other. The procedures also will differ depending on the expectation of further processing by the agency. A procedure for one situation is sometimes quite inappropriate for another.

When the panelists are in conversational communication with each other, they can rely on their natural language for the expression of preferences and the resolution of differences. Otherwise they must rely heavily on abstracted communication such as rating scales, averages, and weights. It should be noted that these abstractions may facilitate good judgment but they do not stand as superior to good judgment. Good judgment is a personal thing, emanating from personal scrutiny and relating to personal experience. Good measurement assists the sharing of experience but is not in itself good judgment. As a group or separately, the panelists seek to make accurate observations but are primarily responsible for tempered judgments.

In the conversational situation, personalities are going to be partly responsible for final choices. Popular and persuasive members will be heeded more than others. The group should be encouraged to reject the idea that unanimity is more important than the thorough review of alternatives. A secret ballot for the final decision may be necessary to give all positions their due.

One of the early responsibilities for the agency or the panel is to identify critical characteristics of the products. A checklist or rating sheet may be useful. Each product needs to be considered, its common features and uniquenesses noted. Logs may be useful.

The panelists should examine the individual products, make such notes to themselves as appropriate, and should meet as a group to discuss the merits of each product. In order to keep their <u>personal</u> judgment scales from stretching and shifting they should consider products simultaneously as well as individually. Relative ordering as well as absolute judging should be a part of the operation, regardless of the form of the final report.

There are two primary approaches, an analytic and a wholistic approach. In the analytic approach each of the important characteristics (attributes) would be given a rating and--by some agreed-upon weighting--an overall rating would be obtained. In the wholistic approach the particular characteristics would be considered but only as subordinate to

the overall rating of merit, which each panelist would provide. Given the usual discrepancies between panel ratings, the distribution of ratings might be registered, or an average might be calculated, or--in the wholistic vein--the panel might continue its discussion until a group score for the panel is agreed upon. This wholistic procedure has the advantage of being less sensitive to oversimplification of product characteristics and less dependent on a proper assignment of weights in the formula. This procedure benefits by focusing panelist attention on his principal task, the delivery of an overall judgment, rather than on an intermediate or alternate task, the judging of characteristics. Sometimes, of course, the product is so complex or diffuse that panelists feel able only to deal with its components. In that instance, of course, a more analytic approach is preferred.

In many applied social science situations, an overly restricted use has been made of the analytic approach. Weighted averages have been used promiscuously. There is the false but common supposition that human judgment does and should operate compensatorily. That is to say, overall judgments are expected as if they came from some weighted combination of component judgments. They are called compensatory because any deficit in one characteristic can be offset by abundance in another.

Obviously people often do not behave that way in handling their personal and institutional affairs. Some deficits are seen as pre-emptive. Some assets are seen as pre-emptive. Any deficiency can disqualify; compensation is not allowed. Some judges seem to respond to a salience that is not ameliorated by other characteristics. Such a person's reasoning is not adequately simulated by prevailing analytic devices. Furthermore, to ask him for an estimate of his weights and to combine them in a weighted sum to estimate overall value usurps his responsibility of providing his best judgment of the overall quality of the product. The problem is alleviated by encouraging the judge to arrive at his terminal judgments in his own way, not restricting him to a particular way of deriving terminal judgments from intermediate judgments or from observations.

Only some people generate terminal ratings that are not reproducible with the usual formulas from their component ratings. Others appear to figure their overall ratings from their own component ratings or make judgments that are consistent with an arithmetic summary. To some outside observers, it is only these latter judges who are rational and trustworthy. The quality of thinking of some of the wholistic judges is too high, however, to disqualify them just because they do not rely heavily on weighted-aggregate conclusions.

The second matter had to do with use of results. It is important for the panelists to know whether or not there may be subsequent processing of their recommendations of merit. If there is, they need to leave some record so that their successors may understand their judgments. For panels operating more analytically, the ratings of components may suffice. For panels operating more wholistically, the dialogue should be abstracted. Here the worksheets used by the panelists may be misleading.

The panel should try to anticipate the agency's actions. There will often be "new conditions" when the agency gets around to using the recommendations. The original recommendations may be much more useful if accompanied by conditional statements. Such statements as these are typical:

1. Product X is rated superior if test-like products remain in contention.
2. Product X is rated superior only if cost is relatively unimportant.
3. Product X is rated superior only if Product A is disqualified.

The agency may not be aware of some conditions that later become important so the panel should identify bases for conditional ratings of products. The ease of dealing with conditional ratings is one of the important advantages of having the panel meeting together. At the outset of any such panel project the agency is in doubt as to how much to set the rules for panel operation and how much to leave them to the panel. If the panel meets as a group, it will want to set its own procedures. It is very likely to find fault with forms and procedures provided for it. The panel may appreciate knowledge of previous panel operations, but its members seem to have a great confidence that it can improve on them. A great deal of time may be taken by a new panel, without appreciably improving on a previous effort. However, the only other way may be to involve the panel in a training session that would take as much time, and still may leave the panel members with feelings that they should have run things their own way.

It is pretty clear that the simpler the forms, the more likely they will be useful to a new panel. Its members should be encouraged not to reject the forms as too simple, but to accept them as way-stations to more intricate considerations. The orientation to a complex form usually will take too long. Hostile feelings often result. The agency should acquaint the new panel with the legacy of previous panels, urge it to set its own procedures, and monitor its progress so that it does not become hooked on the pleasures of creating judging procedures and negligent to the job at hand.